The Dissertation Committee for Sunjoo Hwang
certifies that this is the approved version of the following dissertation:

Three Essays on Contract Theory and Applications

Committee:

Maxwell Stinchcombe, Supervisor

Thomas Wiseman, Supervisor

Fernando Anjos

Venkataraman Bhaskar

Caroline Thomas

Three Essays on Contract Theory and Applications

by

Sunjoo Hwang, B. Eco.; B. Man., M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2015

Soli Deo Gloria

Acknowledgments

I wish to thank my supervisors, Maxwell Stinchcombe and Thomas Wiseman, for their guidance and support, and the members of my dissertation committee, Venkataraman Bhaskar, Fernando Anjos and Caroline Thomas, for their time and valuable advices. I would also like to thank Professors Jaehong Kim, Sawoong Kang and Sonku Kim for opening my eyes into the field of game theory and contract theory. I am indebted to my family for their prayer, support and love. I am very thankful to Professor Byung-Yeon Kim, Professor Sukjin Han, Seung Rae Lee, Donghyun Kim, Daehyun Kim, Sungwon Lee and Jinseok Shin for their support and friendship in graduate school. I would also like to show my gratitude to Kyungmin An for her continuous support, encouragement and prayer during the job market process and the completion of this dissertation.

Three Essays on Contract Theory and Applications

Publication No. ___________________

Sunjoo Hwang, Ph.D.
The University of Texas at Austin, 2015

Supervisors: Maxwell Stinchcombe
Thomas Wiseman

This dissertation consists of three essays. The first essay examines a general theory of information based on informal contracting. The measurement problem—the disparity of true and measured performances—is at the core of many failures in incentive systems. Informal contracting can be a potential solution since, unlike in formal contracting, it can utilize a lot of qualitative and informative signals. However, informal contracting must be self-enforced. Given this trade-off between informativeness and self-enforcement, I show that a new source of statistical information is economically valuable in informal contracting if and only if it is sufficiently informative that it refines the existing pass/fail criterion. I also find that a new information is more likely valuable, as the stock of existing information is large. This information theory has implications on the measurement problem, a puzzle of relative performance evaluation and human resources management. I also provide a methodological contribution. For tractable analysis, the first-order approach (FOA) should be employed. Existing FOA-justifying conditions (e.g. the Mirrlees-Rogerson condition) are so strong that the information ranking condition can be applied

only to a small set of information structures. Instead, I find a weak FOA-justifying condition, which holds in many prominent examples (with multivariate normal or some of univariate exponential family distributions).

The second essay analyzes the effectiveness of managerial punishments in mitigating moral hazard problem of government bailouts. Government bailouts of systemically important financial or industrial firms are necessary ex-post but cause moral hazard ex-ante. A seemingly perfect solution to this time-inconsistency problem is saving a firm while punishing its manager. I show that this idea does not necessarily work if ownership and management are separated. In this case, the shareholder(s) of the firm has to motivate the manager by using incentive contracts. Managerial punishments (such as Obama's $500,000 bonus cap) could distort the incentive-contracting program. The shareholder's ability to motivate the manager could then be reduced and thereby moral hazard could be exacerbated depending on corporate governance structures and punishment measures, which means the likelihood of future bailouts increases. As an alternative, I discuss the effectiveness of shareholder punishments.

The third essay analyzes how education affect workers' career-concerns. A person's life consists of two important stages: the first stage as a student and the second stage as a worker. In order to address how a person chooses an education-career path, I examine an integrated model of education and career-concerns. In the first part, I analyze the welfare effect of education. In Spence's job market signaling model, education as a sorting device improves efficiency by mitigating the lemon market problem. In my integrated model, by contrast, education as a sorting device can be detrimental to social welfare, as it eliminates the work incentive generated by career-concerns. In this regard, I suggest scholarship programs aimed at building human capital rather than sorting students. The second part provides a new perspective on education: education is job-risk hedging device (as well as human capital enhancing or sorting device). I show that highly risk-averse people take high education in order to hedge job-risk and pursue safe but medium-return work path. In contrast, lowly risk-averse people take low education, bear job-risk, and pursue high-risk high-return work path. This explains why some people finish college early and begin start-ups, whereas others take master's or Ph.D. degrees and find safe but stable jobs.

Table of Contents

Chapter 1

Relational Contracts and the Value of Information

1.1 Introduction

The measurement problem is at the core of many failures in incentive system. Murphy (2009) observes that the disparity between true performance and objective performance measures is a cause of the recent 2008-09 financial crisis. Washington Mutual, a large financial institution, rewarded its brokers for their mortgage loan sales. Importantly, bonus payments were based on the quantity rather than the quality of loan sales. Employees thus sold a large amount of bad loans without putting much effort to assess debtors' financial viability of repayment. Finally, the company collapsed during the financial crisis. Countrywide Finance, Wachovia and other smaller lenders experienced similar events. Relatedly, Benabou and Tirole (2013) show that an increasing reliance of Wall Street financial companies on performance pay shifts effort away from less easily contractible but value-enhancing tasks such as long-term investments and risk management. Also, improper asset transformation can arise due to the measurement problem. Myers and Rajan (1998) argue that managers can transform working capital and take the cash, and thereby the inventory is not enough for future sales. This is because "reinvestment on inventory" is hardly contractible. Ample evidence witnesses dysfunctional behaviors or 'gaming' of incentive schemes when performance measures and organizational goals diverge. Heckman et al. (1997) find that government job training agencies "cream skim" by recruiting the easily-placeable candidates rather than the most needy when their rewards are based on job placement of candidates. Similarly, earnings management or stock price manipulation is an important concern when earnings or stock price is primary measure of performance pay. Pourciau (1993) shows empirically that incoming executives manage accruals in an attempt to decrease earnings on the year they come and increase earnings on the following years.

Subjective performance evaluation (or informal incentive contracts) could be a solution to the measurement problem. This is because informal contracts can utilize a lot of information. In order to use an informative signal in a formal contract, it must be verifiable, consistent with social or business norms and incurring little or no cost of writing legal contracts. These many strong restrictions do not apply if a signal is used in an informal manner. Creativity, sincerity, vigilance, within-firm cooperation and other informative signals can be used only in informal contracts but not in formal contracts. In this reason, informal contracts can utilize not just economic but also psychological, social, or even personal information. In the example of Washington Mutual above, firms should have used subjective information such as whether brokers properly consider housing price movement in assessing a loan buyer's financial capability, whether brokers are sincere enough that they do not manipulate objective measures, whether brokers put enough efforts to extract hidden information about customers' financial ability, etc. Likewise, a middle manager's performance can be assessed well if senior management uses various subjective factors such as the middle manager's leadership, initiative and contribution to team members' morale or teamwork. Similarly, designers, white-collar workers, government officials, professors and many other economic agents are informally evaluated based on a large number of qualitative performance measures. Using this large set of informative signals, subjective performance evaluation can better approximate true performance.

However, a disadvantage of informal contracts is that they cannot be enforced by the court but by the contracting parties themselves (i.e. self-enforcement). To motivate an agent, a principal may promise to pay a bonus if the agent's performance—measured by subjective performance signals—is good enough. This promise is non-credible if the principal-agent relationship occurs only once (or finitely many), since the court cannot verify the subjective performance signals and hence cannot enforce the promise. If the relationship repeats over time and is valuable enough, the principal will keep her promise voluntarily in order to secure the valuable relationship. That is, it is not the court but the principal-agent relationship itself that enforces informal contracts. This is the idea of self-enforcing relational contracts. (For this reason, in the following, I will interchangeably use 'subjective performance evaluation' and 'relational contract').

Given this self-enforcement constraint, how informal contracts utilize the large set of information? When and how an additional source information contributes to mitigate the measurement problem? How information affects incentives, efficiency and welfare of contracting parties? The existing literature focuses only on the self-enforcement aspect assuming that only one signal is used but not explores the ramifications of the fact that there is a large set of available signals.

To address these questions, this paper provides a general theory of information in informal incentive contracting. In particular, I find a new *informativeness principle*, a necessary and sufficient condition that characterizes when an additional piece of information is valuable. The principle states that not every statistical information is economically valuable; it is valuable if and only if it is sufficiently informative that it refines existing pass/fail criteria in evaluation. (The difference of my informativeness principle from Holmstrom's standard result will be discussed momentarily). This refinement reduces the disparity of multi-dimensional subjective performance evaluation and true performance. The agent thus faces greater work incentive, and thereby the principal-agent relationship generates greater surplus. Recall that the self-enforcement constraint limits bonus by the total surplus of the principal-agent relationship. As the new information increases the total surplus, the principal can raise the bonus amount, which ends up further increasing the agent's work incentive. That is, the value of information comes from its capability to relax the self-enforcement constraint.

My informativeness principle has following three related theoretic implications. First, it is different from Holmstrom's standard result. Holmstrom (1979) shows that every new source of information is economically valuable in the standard formal contract model, where the main friction is the tradeoff of incentive and risk-sharing, no matter how noisy the information is. In contrast, I show that new source of information is economically valuable in informal contracting environment, where the tradeoff of incentives and self-enforcement is the main friction, if and only if it is decisively informative that the existing evaluation criterion is refined. This divergence sheds new light on a puzzle of relative performance evaluation (RPE). Holmstrom's principle predicts that RPE is superior to absolute performance evaluation (APE), as RPE uses both the agent's and his peer's performance measures in incentive contracting while APE uses only the agent's own performance measures. However, the past three decades of empirical research finds no strong evidence that supports the use of RPE in any formal manner. (See Murphy (1999) and Lewellen (2013)). Relatedly, Gibbons and Murphy (1990) find empirical results that strongly suggest that RPE is used informally but not formally. Holmstrom's result cannot be compromised with their empirical finding since his result is based on formal contracts. My informativeness principle implies that RPE is not necessarily superior to APE in informal contracting. It does if the peer performance measures are sufficiently informative that they refine existing evaluation criterion and hence improve the value of relationship, but not otherwise.

Second, the more a principal already knows about an agent, the better the principal gains from new information. That is, there are *increasing returns to the scale of information*. This property implies that the value of relationship is increasing in members' social, occupational, or demographic similarity. For instance, when a CEO appoints a middle manager of a team of designers, a good middle manager is a designer rather than an accountant or an engineer. This is because the manager-designer accumulates knowledge about designing, and hence, knows how to combine new information about worker-designers with existing knowledge in an attempt to refine evaluation. Similarly, a good manager of a team whose task is firm-specific is an insider of the firm rather than an outsider. This property also implies that (informal) network is a valuable asset, since information is gathered through networking, and hence, dependent board of directors or relational lending could be justifiable: Adams and Ferreira (2007) show that CEO-friendly board of directors, who may be in relational contracting with CEOs, can perform better than independent board, if the advisory role of board is more important than the monitoring role, since the former board members acquire more information from CEOs and hence can give better advices. Hochberg et al. (2007) show that the investment performance of venture capitalists is better if they are in a closer network than in an arm's length so that they could transfer information more effectively.

Third, the second implication above implies that if there is a sufficient stock of existing information, then the disparity between my and Holmstrom's informativeness principles disappears. In this case, one can keep improving the efficiency of informal contracting by adding new source of information repeatedly. But then how far can one improve? I find an *informational folk theorem*, which shows that repeated addition of independent and identical signals to informal contracts leads to the first-best outcome. That is, an optimal subjective performance evaluation is a very large information device, which depends on a myriad of signals. Relatedly, Baker et al. (1994) argue that "the effectiveness of incentive contracting in organizations depends on a large set of social, psychological and economic factors." Given that there is a large set of feasible signals for subjective performance evaluation, unlike it is for objective performance evaluation, the informational folk theorem implies that subjective performance evaluation is a solution to the measurement problem. In this sense of using subjective signals, my information theory based on informal contracting can be a compromise of two seemingly inconsistent theories, informativeness principle and the incomplete contract theory. According to Holmstrom (1979), optimal contracts must depend on every informative signal. However, real-world contracts are often incomplete in the sense that they do not contain terms or conditions for some important contingencies. Existing studies show that it may be optimal to ignore some informative signals if players are concerned with multitasking (Holmstrom and Milgrom (1991)), ambiguity aversion (Mukerji (1998)) or the prohibitive cost of formal contracting. My information theory, by contrast, implies that these signals seem to be ignored but are actually used informally.

In the end, an optimal incentive system is very complicated. Official performance pay (i.e. formal contract) takes only a tiny little part of the whole incentive system. A principal considers a numerous things together in subjective manners to incentivize an agent. This point is related to a recent public sentiment about government regulation on compensation schemes. After the recent 2008-09 financial crisis, the public, market participants and political leaders lost their confidence on the financial sector and thus demanded transparency in managerial compensation schemes. Subjective performance evaluation was discouraged or prohibited. (See United States Troubled Asset Relief Program, which puts objective-measure based restrictions on compensation schemes of the financial companies participating in the program). But this movement goes in a wrong direction (see Murphy (2009)). The use of valuable but subjective information should be encouraged rather than discouraged in order to mitigate the notorious measurement problem, which was a cause of financial crisis. Relatedly, Brunnermeier et al. (2009) are concerned of unforeseen adverse effects of direct regulation on compensation schemes.

The remainder of this paper considers various extensions and a methodological contribution. Hitherto, I compare inclusive information structures. In the first extension, I also compare arbitrary general information structures. I show that an information structure is strictly more valuable than another if (i) the likelihood ratio of the former is a mean-preserving spread (MPS) of the likelihood ratio of the latter and (ii) the signs of the two likelihood ratios are different in any positive measure set.[1] The condition (i) implies that the former information structure contains more informational content than the other, as variations in the likelihood ratio measure the amount of information. Relatedly in formal contracting, Kim (1995) shows that (i) is sufficient for general information ranking. The condition (ii) implies that the pass/fail criteria of evaluation differ across the two information structures. Thus, (i) and (ii) mean that relational contracting requires stronger informativeness conditions than formal contracting does due to the self-enforcement constraint.

The second extension considers general private information structure ranking. Until now I assume that signals are publicly observable by principal and agent. If a performance measure is highly subjective, however, only Principal can observe the performance. This model is an example of the repeated game with private monitoring, which is in general intractable because of its typical lack of recursive form. According to Levin (2003) and Fuchs (2007), I confine my focus to "every-T-period review contracts," a special but realistic set of contracts in which a principal evaluates an agent in every T period. An important difference from the public signal model is that there must be an endogenous termination of the relationship in order to give the truth-telling incentive to the principal and the work incentive to the agent simultaneously. See MacLeod (2003), Levin (2003) and Fuchs (2007). I first characterize optimal evaluation structures and then show that MPS ordering is still valid. I also discuss how the quality of information affects the probability of termination and the sustainability of the relationship.

[1] Let $f(x|a)$ be a probability density function of a vector signal x given a scalar parameter a. $\frac{\partial}{\partial a} \log f(x|a)$ is called the likelihood ratio of the information structure $f(x|a)$.

Finally, I provide a methodological contribution to the relational contract literature. Whether formal or informal, a contract design problem is essentially a constrained maximization problem subject to another maximization problem, and hence potentially very complicated. A principal chooses a contract subject to the incentive-compatibility (IC) constraint in which the agent chooses action. Since the IC constraint is not an (in)equality constraint but an arbitrary maximization constraint, standard techniques like Kuhn-Tucker method cannot be used. Instead, one might suggest the first-order approach (FOA), which replaces the agent's IC constraint with his first-order condition (FOC). This approach seems easily justifiable, as one can assume that the agent's objective function is concave in his action so that IC and FOC are equivalent. However, it turns out that it is a challenge to justify the approach since the agent's objective function involves with a contract, which is an endogenous variable. In formal contract theory, the literature finds some conditions on information structures, such as the Mirrlees-Rogerson condition, that justify the first-order approach. However, these existing FOA-justifying conditions are notoriously restrictive that they are consistent with none of well-known distributions including normal or exponential family. This is problematic because this paper's information ranking criteria are obtained under the assumption that FOA is valid, and hence the ranking criteria can be applied only for a very small set of information structures. In relational contract theory, I find a new FOA justifying condition. My condition is so weak that it is consistent with many of well-known distributions such as normal and exponential family. That is, I provide a first justification of the first-order approach based on relational contract theory.[2] Due to the weakness of my

[2]Kvaloy and Olsen (2014) find a similar justification independently and simultaneously

FOA-justifying condition, my information ranking criteria can be applied to a large set of information structures.

The organization of this paper is as follows. In Section 1.2, I lay out the model. Section 1.3 provides the informativeness principle, the property of increasing returns to the scale of information and the informational folk theorem. As extensions, Sections 1.4 and 1.5 find general public and private information ranking criteria, respectively. In Section 1.6, I justify the first-order approach. I conclude in Section 1.7.

1.2 The Model

In the following, I develop a multi-signal and multitask model of relational contracts. There are two risk-neutral players, Principal (she) and Agent (he) with a common discount factor $\delta \in (0,1)$. In each period $t \in \{1,2,...\}$, they play a stage game in the following way.

Agent works for Principal and contributes to the true output $q_t \in [0,\infty)$. The true output is not (third-party) verifiable (e.g. long-term firm-value) and hence not (formally) contractable. Furthermore, I assume that formal incentive contracting is not a viable option because verifiable signals are either very uninformative or incurring prohibitive verification or legal costs. Principal instead offers Agent a long-term informal contract, which depends on a history of a vector of verifiable and/or nonverifiable signals $x_t \in X \subset \mathbb{R}^n$. x_t may or may not include q_t. If Agent accepts the contract, he chooses a two-dimensional action $a_t = (a_t^1, a_t^2) \in [0,\infty) \times [0,\infty)$. One might interpret a_t^1 as a costly but desirable action, which contributes to the true output q_t, while a_t^2 as

in a specialized setting. The similarity and difference are discussed in section 1.6.

a cheap but undesirable action, which has either no effect or negative effect on the true output. For instance, a_t^1 is an effort to improve long-term firm-value, while a_t^2 is an action of manipulating earnings or stock price. The cost of the desirable action $c(a_t^1)$ satisfies the following usual properties: $c(0) = 0$, $c' > 0$, $c'' < 0$, $\lim_{a_t^1 \to 0} c'(a_t^1) = 0$, and $\lim_{a_t^1 \to \infty} c'(a_t^1) = \infty$. The cost of the undesirable action is $\alpha c(a_t^2)$, where $\alpha \in [0, 1)$. After the choice of action $a_t = (a_t^1, a_t^2)$, the true output q_t and the multi-dimensional signal x_t are realized according to distribution functions $\pi(q_t|a_t)$ and $f(x_t|a_t)$. Let π_1 and π_2 be partial derivatives of π with respect to a_t^1 and a_t^2, respectively. Similarly, I define f_1 and f_2 as partial derivatives of f. After the realization of x_t, either Principal or Agent can choose whether to honor the informal contract.

Basically, the whole contracting problem is a repeated game with imperfect public monitoring with complicated structure. However, Levin (2003) shows that the repeated contracting problem can be greatly simplified to a static contracting problem if the signal and action are both one-dimensional, as for any non-stationary solution to the repeated contracting problem there exists a payoff-equivalent stationary solution. The stationarity result still holds with multi-dimensional signals and actions. See Levin (2003) for detail.

Therefore, I consider only stationary contracts without loss of generality. (I suppress the time-subscript). Let $w \in \mathbb{R}$ be court-enforceable fixed wage and $b(x) \in \mathbb{R}$ be an informally agreed schedule of bonus. Let $\bar{u}^P$ and $\bar{u}^A$ be Principal and Agent's per-period reservation payoffs, respectively. Let $\mathbb{E}[\cdot|a]$ denote conditional expectation. The stationary contracting program is given by

$$\max_{a=(a_1,a_2),w,b(x)} \quad u^P \equiv \mathbb{E}[q|a] - \mathbb{E}[w + b(x)|a] \quad s.t.$$

$$u^P \geq \bar{u}^P \tag{1.1}$$

$$u^A \equiv \mathbb{E}[w + b(x)|a] - c(a_1) - \alpha c(a_2) \geq \bar{u}^A \tag{1.2}$$

$$a \in \arg\max_{a'=(a_1',a_2')} \{\mathbb{E}[w + b(x)|a'] - c(a_1') - \alpha c(a_2')\} \tag{IC}$$

$$b(x) \leq \frac{\delta}{1-\delta}[u^P - \bar{u}^P] \ , \qquad -b(x) \leq \frac{\delta}{1-\delta}[u^A - \bar{u}^A] \quad \forall x \tag{1.3}$$

(1.1) and (1.2) are Principal and Agent's participation constraints. (IC) is Agent's incentive-compatibility constraint, which states that an action a can be induced from Agent if it maximizes Agent's payoff. (1.3) is the condition that ensures informal contracts to be self-enforced: If Principal reneges on bonus payment, Agent leaves. The first inequality of (1.3) thus states that Principal's short-term gain from reneging on bonus payment is smaller than her long-term loss from relationship break-up. The second inequality similarly describes Agent has no incentive to renege on (negative) bonus payment.

Suppose (1.2) is slack in an optimum. Reduce w slightly so that (1.2) is not violated. Then, u^P is increased; (IC) is unaffected; (1.3) is still satisfied by reducing $b(x)$ appropriately for all x. This is a contradiction. Thus, the binding condition of (1.2) characterizes the optimal fixed wage:

$$w = \bar{u}^A + c(a_1) + \alpha c(a_2) - \mathbb{E}[b(x)|a] \tag{BIR}$$

Then, u^P equals $\int q\pi(q|a)dx - c(a_1) - \alpha c(a_2) - \bar{u}^A$. Let $s(a)$ denote the (per-period) total surplus, that is,

$$s(a) \equiv \int q\pi(q|a)dx - c(a_1) - \alpha c(a_2)$$

I assume that $s(a)$ is strictly concave and there is a unique first-best action $a^* = (a_1^*, a_2^*) > 0$ that maximizes $s(a)$. Let $\bar{s}$ be the sum of two players' reservation payoffs $\bar{u}^P$ and $\bar{u}^A$. Then, (1.3) is reduced to the following self-enforcement condition:

$$0 \leq b(x) \leq \frac{\delta}{1-\delta}[s(a) - \bar{s}] \quad \forall x \tag{SE}$$

The last right-hand side term in (SE) is important. If $s(a) \geq \bar{s}$, it is the sum of discounted (per-period) net surpluses. Accordingly, the value of relationship $V(a)$ is defined as

$$V(a) \equiv \max\left\{\frac{\delta}{1-\delta}[s(a) - \bar{s}], 0\right\} \tag{1.4}$$

Finally, the stationary contracting program is further simplified to the following surplus maximization problem:

$$\max_{a,w,b(x)} \; s(a) \quad \text{subject to (BIR), (SE) and (IC)} \tag{P1}$$

Note that w does not play any role in the program above; $s(a)$, (SE) and (IC) are independent of w. Hence, I ignore w for the following analysis. (After finding optimal $b(x)$ and a, w is characterized by (BIR)). Note that (P1) is a maximization program subject to another maximization program (IC) and is typically intractable. The first-order approach (FOA) suggests replacing (IC) with the following first-order condition:

$$\begin{aligned} \int b(x) f_1(x|a)dx &= Cov\left(b(x), \frac{f_1(x|a)}{f(x|a)} \,\middle|\, a\right) = c'(a_1) \\ \int b(x) f_2(x|a)dx &= Cov\left(b(x), \frac{f_2(x|a)}{f(x|a)} \,\middle|\, a\right) = \alpha c'(a_2) \end{aligned} \tag{FOC}$$

where $Cov(\cdot|\cdot)$ is conditional covariance.[3] That is, I solve the following *relaxed*

[3]Dewatripont et al. (1999) originally use this covariance representation in a career concerns model: since $\mathbb{E}\left[\frac{f_a(x|a)}{f(x|a)} \mid a\right] = 0$, it follows that $Cov\left(b(x), \frac{f_a(x|a)}{f(x|a)} \mid a\right) = \int b(x)\frac{f_a(x|a)}{f(x|a)} f(x|a)dx$.

problem

$$\max_{a,b(x)} \; s(a) \quad \text{subject to (SE) and (FOC)} \tag{P2}$$

Note that (FOC) is weaker than (IC). Thus, in general, (P2) has a greater choice set than (P1), and hence, a solution to (P2) is not a solution to (P1). That is, the first-order approach (FOA) is not necessarily valid. One might simply assume that Agent's objective function is concave in action so that (FOC) and (IC) are equivalent. This is invalid argument, however, because Agent's objective function is involved with an endogenous choice variable $b(x)$, for which one cannot *a priori* make an assumption. The seminal work by Mirrlees (1976, 1999) demonstrates that justifying FOA is neither easy nor insignificant, and hence, there is a large body of related literature. Furthermore, to the best of my knowledge, FOA is not yet justified in relational contract framework. Section 1.6 provides a formal justification of FOA and its relevance to the information theory of this paper.

Given that FOA is valid, the following proposition characterizes an optimal informal contract $b(x)$.

Proposition 1.1. *Let μ_1 and μ_2 denote Lagrange multipliers of the first and the second equalities of (FOC), respectively. Let μ denote $(\mu_1, \mu_2)'$. Let $\theta(x|a)$ denote the angle between two vectors μ and $\nabla f(x|a) \equiv (f_1(x|a), f_2(x|a))'$. Suppose that there exists a solution $(a^f, b^f(x))$ to (P2) such that $a^f > 0$. Then, (i) if $\mu \neq 0$, $b^f(x)$ is characterized via*

$$b^f(x) = \begin{cases} 0 & \text{if} \;\; \cos\theta(x|a^f) < 0 \\ \in \left[0, V(a^f)\right] & \text{if} \;\; \cos\theta(x|a^f) = 0 \\ V(a^f) & \text{if} \;\; \cos\theta(x|a^f) > 0 \end{cases} \tag{1.5}$$

where $\cos\theta(x) = \frac{\mu\cdot\nabla f(x|a)}{\|\mu\|\|\nabla f(x|a)\|}$. If $\mu = 0$, then $a^f = a^*$ and $b^f(x)$ is indeterminate: $b^f(x)$ is any function that satisfies (FOC) and (SE) at a^*. (ii) If FOA is *valid, then* $(a^f, b^f(x), w^f)$, *where* w^f *is given by (BIR), is a solution to (P1).*

Proof: Note that $a = a^f > 0$ implies $V(a) > 0$. Suppose not. If $V(a) < 0$, (SE) implies that there is no solution to (P2), a contradiction. If $V(a) = 0$, (SE) implies $b(x) = 0$ for all x. (FOC) then implies $a_1 = a_2 = 0$, a contradiction.

Let $L(x)$ and $G(x)$ denote Lagrange multipliers of the first and the second inequalities, respectively, of (SE). The Lagrange equation is given by

$$\begin{aligned}\mathcal{L} =& s(a) + \mu_1\left(\int b(x)f_1(x|a)dx - c'(a_1)\right) + \mu_2\left(\int b(x)f_2(x|a)dx - \alpha c'(a_2)\right)\\ &+ \int L(x)b(x)dx + \int G(x)[V(a) - b(x)]dx\end{aligned} \tag{1.6}$$

Kuhn-Tucker condition with respect to $b(x)$ is given by

$$\mu_1 f_1(x|a) + \mu_2 f_2(x|a) = G(x) - L(x) \quad \forall x \tag{1.7}$$

By (1.7), $b(x) = V(a)$ if $\mu_1 f_1(x|a) + \mu_2 f_2(x|a) > 0$ since then $G(x) > L(x) = 0$. (Note that it is impossible to have $G(x)$ and $L(x)$ are both positive since then $b(x) = V(a) = 0$, a contradiction). Similarly, $b(x) = 0$ if $\mu_1 f_1(x|a) + \mu_2 f_2(x|a) < 0$. $b(x)$ is indeterminate if $\mu_1 f_1(x|a) + \mu_2 f_2(x|a) = 0$. Thus, the sign of $\mu_1 f_1(x|a) + \mu_2 f_2(x|a)$ determines the optimal bonus $b(x)$. Let μ denote $(\mu_1, \mu_2)'$. Then, it follows that

$$\mu_1 f_1(x|a) + \mu_2 f_2(x|a) = \mu \cdot \nabla f(x|a) = \|\mu\| \|\nabla f(x|a)\| \cos\theta(x|a)$$

Thus, the sign of $\mu_1 f_1(x|a) + \mu_2 f_2(x|a)$ is equivalent to the sign of $\cos\theta(x|a)$.

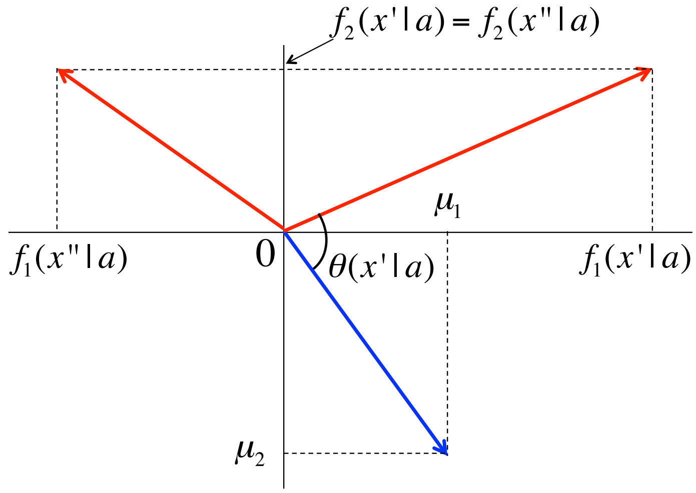
Suppose that $\mu_1 = \mu_2 = 0$. Then Kuhn-Tucker conditions with respect to a_1 and a_2 imply that $s_1(a) = s_2(a) = 0$, which means $a = a^*$. In this case,

$\cos\theta(x|a^*) = 0$. So any function $b^f(x)$ that satisfies (FOC) and (SE) at a^* is optimal. ■

Proposition 1.1 states that optimal relational contract $b^f(x)$ consists of two factors, bonus and target. By the self-enforcement constraint (SE), incentives are bounded below by the value of relationship $V(a^f)$. Since the players are risk-neutral, and hence they have linear preferences, incentives need not be too much dependent on signal realization: paying the maximum bonus $V(a^f)$ if x falls in a "pass" region and paying no bonus if x falls in a "fail" region is optimal. The pass/fail criterion is solely dependent on the cosine function $\cos\theta(x|a^f)$, which measures the extent to which Agent's contribution to the measured performance $\nabla f(x|a^f)$ is close to his contribution to the true performance μ. If the contribution to the measured performance and that to the true performance are in similar directions (in the sense that $\cos\theta(x'|a^f) > 0$) at a signal level x', then x' belongs to the pass region. If the two contributions are in different directions (in the sense that $\cos\theta(x''|a^f) < 0$) at x'', then x'' belongs to the fail region. Figure 1.1 illustrates this pass/fail criterion. The vector μ reflects the idea that the action a_1 is desirable and hence contributing to the true output in the sense that $\mu_1 > 0$, whereas the action a_2 is undesirable in the sense that $\mu_2 \leq 0$. Thus, Principal needs to encourage Agent to choose high a_1 but low a_2. Consider two performance levels x' and x''. To make x', Agent should put relatively high a_1 but low a_2 (in the sense that $\frac{f_1(x'|a)}{f_2(x'|a)}$ is sufficiently large). To obtain x'', Agent should put relatively low a_1 but high a_2 (in the sense that $\frac{f_1(x'|a)}{f_2(x'|a)}$ is sufficiently small). Then, it is optimal to compensate Agent if x' is realized while penalize him if x'' is realized.

Relatedly, Gibbons (2010) considers a similar cosine model in a lin-

Figure 1.1: Optimal pass/fail criterion with multitask



ear formal contracting context. A noteworthy point is that the linear pay-performance sensitivity increases in the *size* of cosine, which measures the similarity of contribution to true and that to measured performances. In contrast, the size of cosine in this general relational contract model does not affect the size of incentives (i.e. bonus). Only the sign of cosine matters: it determines the pass/fail criterion given a fixed bonus.

Recall that bonus and target (i.e. the pass/fail criterion) are two components of optimal relational contracts. Note that bonus is fixed (as $V(a)$) and is independent of information structure $f(x|a)$ if a is fixed. In contrast, target can be potentially improved if better information structure is employed. Since we consider informal contracting for which there is a large set of available signals, one can easily find better information structures that adds additional signals. Therefore, for the remainder of this paper, I examine how better information structure contributes to refine the target and hence mitigate measurement problems. To this end, I develop information ranking orders that compare two arbitrary information structures $f(x|a)$ and $h(\cdot|a)$.

To focus on information theory, I assume that the undesirable action a_2 is cheap and useless in the sense that $\alpha = 0$ and $\int q\pi_2(q|a)dq = 0$ for any a. The Kuhn-Tucker condition of (1.6) with respect to a_1 then implies that $\mu_2 = 0$ for any a. If $\mu_2 = 0$, the target is independent of $f_2(x|a)$. Then, Agent will optimally choose $a_2 = 0$. Therefore, I ignore the choice of a_2 in the following. Also I use the following notation

Notation 1.1. $a = (a_1, 0) = a_1$ *and* $f_a(x|a) = f_1(x|a_1, 0)$

In this case, the solution to relational contracting problem becomes simpler and hence we can get to know additional properties about equilibrium action a^f. See the following related corollary.

Corollary 1.1. *Suppose that the first-order approach is valid and that there exists a solution* $(a^f, b^f(x))$ *to (P2) such that* $a^f > 0$. *Then, (i)* $\mu_1 \geq 0$. *(ii)* $a^f \leq a^*$. *(iii) If* $a^f < a^*$, *then* $b^f(x)$ *is characterized via*

$$b^f(x) = V(a^f) 1_{\left\{\frac{f_a(x|a^f)}{f(x|a^f)} > 0\right\}} \tag{1.8}$$

(iv) Suppose $a^f = a^*$. $b^f(x)$ *can be any contract satisfying (SE) and (FOC).*[4]

[4]Note that the bonus contract above is not renegotiation-proof: once either party reneges on the bonus payment, the players find that continuing the relationship is Pareto superior to termination since they can then obtain some positive surplus. However, Levin (2003) provides an easy way to modify the bonus contract so that it is renegotiation-proof (robust to the self-reference argument). Suppose that the parties use the bonus contract on the equilibrium path but instead initiate the following penalty contract $(\hat{w}, \hat{b}(x))$ on the deviation path: $\hat{b}(x) = -V(a^f) 1_{\left\{\frac{f_a(x|a^f)}{f(x|a^f)} \leq 0\right\}}$ and $\hat{w}$ is such that $\mathbb{E}[\hat{w} + \hat{b}(x)|a^f] - c(a^f) = s - \bar{u}^P$. That is, the relationship continues and Agent exerts the same action level a^f even in the deviation path, whereas, in contrast to the equilibrium path, Principal takes only $\bar{u}^P$ and Agent takes $(s - \bar{u}^P)$. However, in this phase, Agent may deviate by reneging on the penalty payment. Then, they move back to the original bonus contract to punish Agent. Thus, this combination of bonus and penalty contracts is renegotiation-proof.

That is, an optimal relational contract has the bang-bang type in (1.8). To provide the highest incentive, evaluation should be as accurate as possible and compensation for good performance should be the largest. In the evaluation, Principal reviews the multi-dimensional information $x = (x_1, ..., x_n)$ to determine Agent's pass/fail, that is, the *sign* of the likelihood ratio $\frac{f_a(x_1,...,x_n|a)}{f(x_1,...,x_n|a)}$. In the compensation, Agent gets paid a single bonus $V(a)$ if he passes the evaluation. In real life, organizations prevalently use the Balanced Score Card (BSC), which lists various targets employees are supposed to meet. For instance, a typical BSC for human resource managers sets (1) hiring and keeping skilled personnel, (2) maintaining safe and legal human resource management, (3) enhancing workers' loyalty to the company and many other goals. Such a complicated review is followed by simple compensation such as promotion or fixed monetary reward. In this sense, I call the bang-bang type contract as the multi-target single-bonus contract.

Levin (2003) attains the same result when there is a single signal $x \in \mathbb{R}$. But I extend his result in two senses. First, I consider a multi-signal $x = (x_1, ..., x_n)$ so that I can examine the value of information with relational contracting. Second, I show that the optimality of the bang-bang type contract holds even under the absence of the following standard but notoriously restrictive assumptions (on information structure) that he uses.

Assumption 1.1. *Monotone likelihood ratio property (MLRP):*
$\frac{f_a(x|a)}{f(x|a)}$ *is strictly increasing in* x *for each* a.[5]

MLRP implies that x is stochastically increasing in a. In this sense, one can understand a as a productive effort.

[5] In contract theory literature, $\frac{\partial}{\partial a} \log f(x|a)$ is often called the likelihood ratio of $f(x|a)$.

Assumption 1.2. *Convexity of distribution function condition (CDFC): $F(x|c^{-1}(\kappa))$ is strictly convex in κ, where $\kappa = c(a)$.*

The absence of the two conditions above (and replacing them with other weak conditions) have following theoretic implications. First, Contract theory literature (formal, informal and incomplete contracting) pervasively assumes MLRP. In this relational contract model, however, I do not need MLRP since the sign of likelihood ratio itself is a sufficient statistic that captures every relevant information. This is a good thing because MLRP is increasingly restrictive as the dimension n of a vector signal $x = (x_1, ..., x_n)$ rises.

Second, the combination of MLRP and CDFC is called the Mirrlees-Rogerson condition, which is a standard condition to justify the first-order approach (FOA) in formal contracting. However, this Mirrlees-Rogerson condition is so restrictive that it is inconsistent with almost every well-known probability distributions (including normal and exponential family). This is problematic in particular in this paper. It is because this paper focuses on finding general information ranking orders that can compare two arbitrary information structures. The Mirrlees-Rogerson condition greatly shrinks the set of information structures for which the ranking orders can be applied since the ranking orders will be obtained by assuming that FOA is valid. In Section 1.6, I find a weak condition that justifies FOA in relational contracting framework. This new FOA-justifying condition is consistent with normal and some of exponential family distributions.

Example 1.1. *The exponential family: Let $x = (x_1, ..., x_n)$ be a vector and a be a scalar. Suppose that $f(x|a)$ is a probability density function that belongs to the exponential family*

$$f(x|a) = h(x) \exp\left(\eta(a)T(x) - \psi(a)\right) \tag{1.9}$$

where $h(x)$, $T(x)$, $\eta(a)$ and $\psi(a)$ are some known functions and I assume that $\eta'(a) \neq 0$. The additive-normal case is an example of this exponential family. $T(x)$ is particularly important in statistics since it is a sufficient statistic of x. That is, the level of $T(x)$ is statistically informative. However, not the level but the (constant-adjusted) sign of $T(x)$ is economically valuable in the relational contract model: note that $\mathbb{E}[T(x)|a] = \frac{\psi'(a)}{\eta'(a)}$ and $\frac{f_a(x|a)}{f(x|a)} = \eta'(a)T(x) - \psi'(a)$. Therefore, $\frac{f_a(x|a)}{f(x|a)} > 0$ if and only if

$$T(x) > \mathbb{E}[T(x)|a] \tag{1.10}$$

The sign of (constant-adjusted) $t = T(x_1, ..., x_n)$ is a sufficient statistic of the total information $x = (x_1, ..., x_n)$. That is, t is a one-dimensional composite performance measure.

1.3 Information Ranking 1: Inclusive (Public) Signals

The previous section shows that optimal informal contracts consist of two parts, bonus and target. Optimal bonus is independent of information structure. But optimal target can be potentially refined if an additional source of information is used. Given the large set of available signals, one can easily find an additional signal. In this section, I examine when and how an additional signal refines optimal target, and thereby improve relational contracting.

There is an existing multi-signal $x \in X \subset \mathbb{R}^{n_1}$ with density $f(x|a)$, $n_1 \in \mathbb{N}$. There is also an additional multi-signal $y \in Y \subset \mathbb{R}^{n_2}, n_2 \in \mathbb{N}$. Let $h(x, y|a)$ be the density of the pair of signals (x, y) such that $f(x|a) = \int h(x, y|a)dy$ for each a. Both f and h are everywhere positive and differentiable in a, with partial derivatives f_a and h_a, respectively. In the following, I compare these two information structures (X, f) and (X, Y, h) in terms of their effects on optimal relational contracts and efficiency.

Clearly, Principal is *weakly* better off using the new signal y in addition to x in making a contract because (X, Y, h) contains at least as much information as (X, f) and Principal has the option to ignore y. To see this, note that $h(x, y|a) = p(y|x, a)f(x|a)$ for some positive conditional density $p(y|x, a)$. Let p_a denotes partial derivative of $p(y|x, a)$. Then, it follows that

$$\frac{h_a(x, y|a)}{h(x, y|a)} = \frac{f_a(x|a)}{f(x|a)} + \frac{p_a(y|x, a)}{p(y|x, a)} \tag{1.11}$$

Note that the variation of the likelihood ratio is an amount of information.[6] Thus, the new signal y provides (degenerate or non-degenerate) additional information, which is measured by the variation in $\frac{p_a}{p}$.

When is Principal strictly better off? In the standard static formal contract model in which Agent is risk-averse, Holmstrom (1979) shows that Principal is strictly better off (making Agent indifferent) if and only if the additional information is non-degenerate. Consider the following formal definition of degeneracy.

Definition 1.1. For two information structures (X, f) and (X, Y, h), x is a sufficient statistic for (x, y) in estimating a if for all a,

$$\frac{h_a(x, y|a)}{h(x, y|a)} = \frac{f_a(x|a)}{f(x|a)} \ \ a.e. \ (x, y) \tag{1.12}$$

[6]For instance, the variance in the likelihood ratio, the Fisher information, is a linear measure of the variation in the likelihood ratio. In a well-defined statistical inference problem, this linear measure is a complete order of information structures. In contract models, however, I often need a more general measure, the mean-preserving spread in likelihood ratios.

Or equivalently, $\frac{p_a(y|x,a)}{p(y|x,a)} = 0$ a.e. (x, y).[7]

If (and only if) x is not a sufficient statistic, the variation in $\frac{p_a}{p}$ is non-degenerate. In this case, (x, y) contains more information than x does, that is, y is statistically informative.[8] Thus, Holmstrom's informativeness principle states that in the standard formal contract model a new signal y is economically valuable if and only if it is statistically informative.

Consider the following related concept "non-sign-sufficiency." I shall show that it is a complete order of information structures in the relational contract model.

Definition 1.2. For two information structures (X, f) and (X, Y, h), x is a sign-sufficient statistic for (x, y) in estimating a, if all of the following three sets are measure-zero

$$\left\{(x, y) : \frac{f_a}{f} > 0, \frac{h_a}{h} < 0\right\}, \left\{(x, y) : \frac{f_a}{f} < 0, \frac{h_a}{h} > 0\right\}, \left\{(x, y) : \frac{f_a}{f} = 0, \frac{h_a}{h} \neq 0\right\} \tag{1.13}$$

for all a, where I suppress arguments of likelihood ratios for notational simplicity.[9]

If x is not a sign-sufficient statistic for (x, y), the new signal y changes the sign of the likelihood ratio in a nondegenerate manner. The following proposition is a main result of this paper, which shows that in the relational

[7]Holmstrom (1979) confines his focus to distributions such that (1.12) holds for all a or no a arguing that other distributions are of little interest. Similarly, I confine my focus to distributions such that (1.13) holds for all a or no a.

[8]Note that $h(x, y|a)$ is more informative than $f(x|a)$ in Blackwell's notion whether x is sufficient or not for (x, y) in estimating a.

[9]I confine my focus to distributions such that (1.13) holds for all a or no a, as Holmstrom (1979) does arguing that other distributions are of little interest.

contract model a new signal y is economically valuable if and only if it is not sign-sufficient (i.e. sign-changing).

Proposition 1.2. *Suppose the first-order approach holds under both* (X, f) *and* (X, Y, h). *Suppose that (P2) has solutions* $(a^f, b^f(x))$ *and* $(a^h, b^h(x, y))$ *under* (X, f) *and* (X, Y, h), *respectively, and that* $0 < a^f < a^*$. *Then,* (X, Y, h) *strictly Pareto improves* (X, f) *if and only if* x *is not a sign-sufficient statistic for* (x, y) *in estimating* a.

Proof: Consider the following two terms $\mathbf{H}(a)$ and $\mathbf{F}(a)$—maximized marginal incentives under information structures (X, Y, h) and (X, f), respectively.

$$\mathbf{H}(a) \equiv Cov\left(V(a)\mathit{1}_{\{\frac{h_a}{h}>0\}}, \frac{h_a}{h}\middle| a\right)$$

$$\mathbf{F}(a) \equiv Cov\left(V(a)\mathit{1}_{\{\frac{f_a}{f}>0\}}, \frac{f_a}{f}\middle| a\right)$$

where I suppress arguments from the likelihood ratios for notational simplicity.

The proof consists of the following steps. Step 1 shows that (X, Y, h) and (X, f) are payoff-equivalent if '$\mathbf{H}(a) = \mathbf{F}(a)$ for all a.' Step 2 shows that (X, Y, h) strictly Pareto improves (X, f) if '$\mathbf{H}(a) > \mathbf{F}(a)$ for all a.' Step 3 shows that '$\mathbf{H}(a) > \mathbf{F}(a)$ for all a' if x is not sign-sufficient. Step 4 shows that '$\mathbf{H}(a) = \mathbf{F}(a)$ for all a' if x is sign-sufficient.

Step 1: Recall that $\mathbf{F}(a)$ is the maximized marginal incentive under (X, f). By the optimality of multi-target single-bonus contract (see Proposition 1.1), a^f is the largest solution to the equation $\mathbf{F}(a) = c'(a)$. Suppose that a' is the largest solution to the equation $\mathbf{H}(a) = c'(a)$. Then, $a' = a^f < a^*$. But then by the optimality of multi-target single-bonus contract (see Proposition 1.1), a^h equals a', which means $a^f = a^h$.

Step 2: The equilibrium payoffs of Principal and Agent are $s(a) - \bar{u}^A$ and $\bar{u}^A$ for any information structure since Agent's participation constraint is binding. Then, (X, Y, h) strictly Pareto improves on (X, f) if and only if $a^h > a^f$, since $s(a)$ is strictly concave, $a^f < a^*$ and $a^h \leq a^*$ by Proposition 1.1. Suppose $a^h = a^*$. Then, strict Pareto improvement immediately follows. Suppose instead $a^h < a^*$. Then, by the optimality of multi-target single-bonus contract (see Proposition 1.1), a^h is the largest solution to the equation that $\mathbf{H}(a) = c'(a)$. If $a^f = a^h$, then $\mathbf{F}(a^f) = c'(a^f) = c'(a^h) = \mathbf{H}(a^h)$, a contradiction to the hypothesis. Therefore, we have (i) $a^f \neq a^h$. Note that under (X, Y, h) Principal can always induce at least as much action as a^f simply by ignoring the new signal y. That is, (ii) $a^h \geq a^f$. (i) and (ii) imply that $a^h > a^f$.

To prove step 3 and 4, note that (1.11) implies the following martingale property:

$$\frac{f_a(x|a)}{f(x|a)} = \mathbb{E}\left[\frac{h_a(x,y|a)}{h(x,y|a)}\,\middle|\,x, a\right] \tag{1.14}$$

(1.14) always holds whether x is sign-sufficient or not. Then, it follows that

$$
\begin{aligned}
\mathbf{F}(a) &= \mathbb{E}\left[V(a)1_{\left\{\frac{f_a}{f}>0\right\}}\frac{f_a}{f}\,\middle|\, a\right] = \mathbb{E}\left[V(a)1_{\left\{\frac{f_a}{f}>0\right\}}\mathbb{E}\left[\frac{h_a}{h}\,\middle|\, x, a\right]\,\middle|\, a\right] \\
&= \mathbb{E}\left[V(a)1_{\left\{\frac{f_a}{f}>0\right\}}\frac{h_a}{h}\,\middle|\, a\right] = V(a)\int_{\left\{\frac{f_a}{f}>0\right\}}\frac{h_a}{h}hd(x,y) \\
&= V(a)\int_{\left\{\frac{f_a}{f}>0,\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) + V(a)\int_{\left\{\frac{f_a}{f}>0,\frac{h_a}{h}<0\right\}}\frac{h_a}{h}hd(x,y) \\
&\le V(a)\int_{\left\{\frac{f_a}{f}>0,\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) \qquad (1.15)\\
&\le V(a)\int_{\left\{\frac{f_a}{f}>0,\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) + V(a)\int_{\left\{\frac{f_a}{f}<0,\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) \\
&\quad + V(a)\int_{\left\{\frac{f_a}{f}=0,\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) \\
&= V(a)\int_{\left\{\frac{h_a}{h}>0\right\}}\frac{h_a}{h}hd(x,y) = \mathbb{E}\left[V(a)1_{\left\{\frac{h_a}{h}>0\right\}}\frac{h_a}{h}\,\middle|\, a\right] = \mathbf{H}(a)
\end{aligned}
$$

Step 3: Suppose x is not sign-sufficient. Then at least one of $\eta_1 \equiv \left\{\frac{f_a}{f} > 0, \frac{h_a}{h} < 0\right\}$, $\eta_2 \equiv \left\{\frac{f_a}{f} < 0, \frac{h_a}{h} > 0\right\}$ and $\eta_3 \equiv \left\{\frac{f_a}{f} = 0, \frac{h_a}{h} \neq 0\right\}$ is of positive measure. If η_3 is of positive measure, then $\left\{\frac{f_a}{f} = 0, \frac{h_a}{h} > 0\right\}$ is also of positive measure by (1.14). Therefore, either one of the two inequalities in (1.15) is strict. Then, $\mathbf{H}(a) > \mathbf{F}(a)$ for all a.

Step 4: Suppose instead that x is sign-sufficient. Then, η_1, η_2 and η_3 are all measure-zero. Hence, both inequalities in (1.15) become equalities. ■

Proposition 1.2 states that a new signal is valuable in the sense of strict Pareto improvement if and only if it is a sign-changing signal. The intuition is as follows. Given the risk-neutrality and the self-enforcement constraint, the multi-target single-bonus contract is optimal (unless the first-best action

a^* is inducible), where the optimal target depends solely on the sign of the likelihood ratio. The addition of a new signal y can potentially refine this target. If y is sufficiently informative, the variation in $\frac{p_a}{p}$ is large enough. In this case, the sign of new likelihood ratio $\frac{h_a}{h} = \frac{f_a}{f} + \frac{p_a}{p}$ is different from the sign of the existing likelihood ratio $\frac{f_a}{f}$ in some regions. Therefore, there is a region in which "pass" is refined to "fail" and another region in which "fail" is refined to "pass."

The refinement in evaluation criterion initiates the following adjustment in contracts. As the refined pass/fail criterion better approximates the desirable action a, Agent faces higher work incentive given a fixed bonus amount. This higher incentive leads to an increase in the value of relationship. Principal can raise the bonus amount, as the limit on the bonus (i.e. the value of relationship) is relaxed. The higher bonus induces even more incentive. The value of relationship increases again, and thus bonus can be increased further, and the like. This is an economics of how a new signal improves existing relational contracts and thereby enhances efficiency.

Furthermore, I can also obtain comparative statics: how a new signal affects action and compensation. Agent expects more compensation, but at the same time exerts more action so that his expected utility is unchanged. See the following corollary.

Corollary 1.2. $a^h > a^f$, $\mathbb{E}[w^h + b^h(x, y)|a^h] > \mathbb{E}[w^f + b^f(x)|a^f]$.

Proof: $a^h > a^f$ immediately follows from the proof of Proposition 1.2. Since (1.2) is binding with any information structure, it follows that $\mathbb{E}[w^i + b^i(\cdot)|a^i] = \bar{u}^A + c(a^i)$ for $i = h, f$. Thus, the expected payment is strictly increasing in the induced action. ■

A virtue of this informativeness principle based on relational contracting is that it provides detailed economics of information. In contrast, in the standard formal contract model, the economics of how new information improves efficiency is a black box: Holmstrom (1979) shows that new information refines target (which depends one-to-one on the level of the likelihood ratio), and thereby leads to strict Pareto improvement. But it is unknown how the new target looks like and how incentives depend on target. Furthermore, it is unknown with formal contracts whether the equilibrium action increases or decreases or whether the expected compensation increases or decreases. This is because a functional form of optimal contracts is known in the relational contract model but not in the standard formal contract model.

"Non-sign-sufficiency" (or sign-change) is a stronger condition than "non-sufficiency" (or level-change), which is equivalent to statistical informativeness. Hence, some informative signals are not valuable in the relational contract model. To see this difference, consider the standard formal contract model. Agent is risk-averse having a concave utility function $u(\cdot)$. An optimal formal contract $W(x)$ is given by the following standard characterization:

$$\frac{1}{u'(W(x))} = \lambda + \mu \frac{f_a(x|a)}{f(x|a)} \tag{1.16}$$

where $\lambda > 0$ and $\mu > 0$. To incentivize Agent, $W(x)$ must be a function of $\frac{f_a}{f}$. Because of the concavity of $u(\cdot)$, $W(x)$ must be a smooth function of $\frac{f_a}{f}$, and hence depends one-to-one on the level of $\frac{f_a}{f}$. Thus, the trade-off between risk and incentive leads Principal to care for not just Agent's pass/fail (i.e., the sign of the likelihood ratio) but also how good/bad (i.e., the level of the likelihood ratio) he is. See Table 1.1.

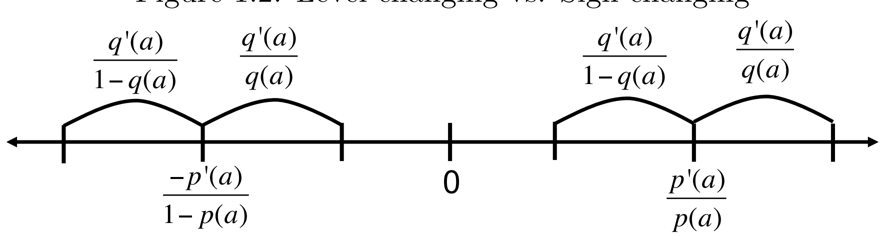
The following example provides a signal, which is level-changing but not sign-changing.

Example 1.2. *Consider two discrete random signals x and y such that $x \in \{1, -1\}$, $y \in \{1, -1\}$ and a continuous action choice $a \in [0, \infty)$. The information structure is given by*

$x,\ y$	1	-1
$\mathbb{P}(x\|a)$	$p(a)$	$1 - p(a)$
$\mathbb{P}(y\|a)$	$q(a)$	$1 - q(a)$

where $p(a), q(a) \in (0,1)$ and $p'(a) > 0$ for each a. (Note that x and y are independent conditional on a). Given x, the addition of y provides an inclusive information, which is measured by the additional variation in the likelihood ratio (see Figure 1.2). x is sufficient for (x, y) in estimating a if and only if $|q'(a)| = 0$. Take a such that $q'(a) \geq 0$. Then, x is sign-sufficient for (x, y) in estimating a if and only if $|q'(a)| \leq p'(a) \min\left\{\frac{q(a)}{1-p(a)}, \frac{1-q(a)}{p(a)}\right\}$. Take instead a such that $q'(a) \leq 0$. Then, x is (sign)-sufficient if and only if $|q'(a)| \leq p'(a) \min\left\{\frac{1-q(a)}{1-p(a)}, \frac{q(a)}{p(a)}\right\}$. That is, the inclusion of y creates value in formal contracting if it has any incremental information about a, that is, $|q'(a)| > 0$, whereas it creates value in relational contracting if it has a sufficient amount of incremental information, that is, $|q'(a)| \gg 0$.

Figure 1.2: Level-changing vs. Sign-changing



Example 1.2 illustrates that there is a new signal, which is statistically informative (i.e. level-changing) but not economically valuable (i.e. not sign-changing) if an existing signal x is binary. This strict difference is (generically) true for any discrete signal x.[10]

In contrast, the following example illustrates that "economic valuableness" and "statistical informativeness" are (generically) equivalent if an existing signal x is continuous.

Example 1.3. *Consider the following bivariate random signals* (x, y)*:*

$$\begin{array}{l} x = \mu_1 a + \varepsilon_1 \\ y = \mu_2 a + \varepsilon_2 \end{array}, \quad \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

This is an additive-normal example. An issue is that the first-order approach (FOA) is not yet formally justified in this case. It is because this additive-normal example dose not satisfy existing FOA-justifying conditions (such as the Mirrlees-Rogerson condition). Fortunately, I find a weak FOA-justifying condition, which is consistent with this additive-normal example. See Section 1.6. The likelihood ratio of the existing signal x *is given by*

$$\frac{f_a(x|a)}{f(x|a)} \equiv \mu_1 \frac{x - \mu_1 a}{\sigma_1^2}$$

And, the likelihood ratio of the pair of signals (x, y) *is given by*

$$\frac{h_a(x, y|a)}{h(x, y|a)} \equiv \mu_1 \frac{x - \mu_1 a}{\sigma_1^2} + \left[\frac{\rho\mu_1\sigma_2 - \mu_2\sigma_1}{(1 - \rho^2)\sigma_1\sigma_2}\right] \left[\frac{\rho(x - \mu_1 a)}{\sigma_1} - \frac{(y - \mu_2 a)}{\sigma_2}\right]$$

Therefore, x *is sufficient for* (x, y) *if and only if*

$$\rho = \frac{\mu_2/\sigma_2}{\mu_1/\sigma_1} \quad \textit{or equivalently,} \quad \frac{\mu_2}{\sigma_2} - \rho\frac{\mu_1}{\sigma_1} = 0 \tag{1.17}$$

[10] Consider a knife-edge case in which a discrete signal x has a realization, say x_0, such that $\frac{\partial}{\partial a}\mathbb{P}(x_0|a)$ is exactly equal to 0 and any new signal y is independent of x conditional on a. In this case, a new signal y is sign-changing whenever it is level-changing.

Note that μ_i/σ_i is the signal-to-noise ratio, which measures the amount of information. (1.17) implies that y provides no incremental information because every information y provides is already contained in x through their correlation. Note that x is sign-sufficient for (x, y) if and only if the same condition (1.17) holds. This is because $\frac{f_a(x|a)}{f(x|a)}$ is continuous, and hence, is dense at around the critical level 0. Then, any additional variation in the likelihood ratio caused by the inclusion of y could change the sign of the likelihood ratio. That is, level-sufficiency and sign-sufficiency are equivalent.

As Example 1.3 illustrates, "non-sign-sufficiency" and "sign-sufficiency" are generically equivalent if an existing signal x is continuous. The following proposition shows that if y is independent of x conditional on a, then the two notions are exactly equivalent.

Proposition 1.3. *Suppose that y is independent of x conditional on a and that $\frac{f_a(x|a)}{f(x|a)}$ is a continuous random variable. Then, x is a sign-sufficient statistic for (x, y) in estimating a if and only if x is a sufficient statistic for (x, y) in estimating a.*

Proof: I shall prove only that "non-sufficiency" implies "non-sign-sufficiency" since the other direction is obviously true. By conditional independence, we have $p(y|x, a) = p(y|a)$. By "non-sufficiency," we have $\mathbb{P}\left(\frac{p_a(y|a)}{p(y|a)} \neq 0\right) > 0$, which implies that $\mathbb{P}\left(\frac{p_a(y|a)}{p(y|a)} < 0\right) > 0$ since $\mathbb{E}\left[\frac{p_a(y|a)}{p(y|a)}\right] = 0$. Then, $\exists\, d < 0$ such that $\mathbb{P}\left(\frac{p_a(y|a)}{p(y|a)} \leq d\right) > 0$. Let $\hat{\eta} \equiv \left\{0 < \frac{f_a(x|a)}{f(x|a)} < \frac{-d}{2}, \frac{p_a(y|a)}{p(y|a)} \leq d\right\}$. It follows that (i) $\hat{\eta} \subset \eta_1 \equiv \left\{\frac{f_a(x|a)}{f(x|a)} > 0, \frac{h_a(x,y|a)}{h(x,y|a)} < 0\right\}$ since $d < 0$ and $\frac{h_a(x,y|a)}{h(x,y|a)} = \frac{f_a(x|a)}{f(x|a)} + \frac{p_a(y|a)}{p(y|a)}$. Note that $\frac{f_a(x|a)}{f(x|a)}$ has a strictly increasing distribution function since it is a continuous random variable. Then, it follows that (ii) $\mathbb{P}(\hat{\eta}) > 0$.

Finally, (i) and (ii) imply that $\mathbb{P}(\eta_1) > 0$, meaning x is not sign-sufficient for (x, y) in estimating a. ■

1.3.1 Implications of the Information Theory Based on Relational Contracting

Informational increasing returns to scale: The information theory based on relational contract (i.e. Propositions 1.2 and 1.3 and Examples 1.2 and 1.3) implies that a new source of statistical information y is more likely economically valuable if the existing information x is fine (i.e. continuous) rather than coarse (i.e. discrete). If Principal has a lot of existing information about Agent, she can understand how to combine a new source of information with the existent information so that she can better assess Agent's performance. In this sense, there is increasing returns (of additional information) to scale: the larger the existing stock of information, the more likely a new source of information is valuable.

Relationship is a valuable asset: The increasing returns to information has following implications on organizational economics. First, in human resources management, CEOs should appoint as a middle manager of a team of workers the one who shares occupational or social background with the workers. For instance, if workers are designers, a good middle manager is a designer rather than an accountant or an engineer. This is because the manager-designer accumulates more knowledge about designing, and hence, knows how to combine new information about worker-designers with the existing knowledge in order to refine evaluation. Similarly, a good middle manager of a team whose task is firm-specific is an insider of the firm rather than an outsider. Second, in corporate governance, CEO-friendly boards of directors could out-perform independent boards, if the advisory role is more important than the monitoring role. Board members who are connected to a CEO through a relational contracting could acquire more information from the CEO and hence give better advice. See Adams and Ferreira (2007). Third, relational lending could perform better than arms-length lending, as lenders share more information within a close network, which contributes to investment performance. See Hochberg et al. (2007).

A compromise to the conflict of informativeness principle and real world contracts: Consider the following related example. Let y be a medical status of Agent's wife. The signal y will be most likely highly informative of Agent's effort a. According to Holmstrom's standard informativeness principle, this personal matter must be incorporated in a formal contract. However, real world contracts very rarely have explicit reference to these nonstandard information. Furthermore, as Murphy (1999) observes, formal contracts are not contingent on relatively clear information such as Agent's peer's performance. Relatedly, Holmstrom (1979) suggests that "note, however, that internal labor contracts rarely contain explicit reference to monitoring information ... Yet such information is and should be used. The reason the principal (i.e. the firm) will not default on such an implicit contract is its concern for reputation in the labor market." Thus, He conjectures that many informative signals are not used in formal contracts but should be used at least in informal contracts. The information theory in this paper shows that his conjecture can be confirmed if these signals are so informative that they can contribute to improve the value of relationship (in the sense of sign-change), but not otherwise. If Principal accumulates a lot of information about coworkers (or agents) through long-term relationship so that statistical information is economically valuable, then an optimal relational contract is a function of a myriad of standard and non-standard signals, as Baker et al. (1994) suggest: "the effectiveness of incentive contracting in organizations depends on a large set of social, psychological and economic factors."

A curse of dimensionality: One might argue that optimal relational contracts are hardly implementable due to the curse of dimensionality since many signals are used, and hence evaluation is complicated. However, the burden of describing multi-dimensional signals $x = (x_1, ..., x_n)$ is smaller than it might first appear to be: Principal needs to consider only the sign of the likelihood ratio of x, a sufficient statistic of the total information x, rather than the level of x. If the sign of the likelihood ratio is still hard to describe, then one can consider a simple but still informative garbling x' of x: let x' be either 1 (satisfactory) or 0 (unsatisfactory). (Similarly, one might consider more fine evaluation such as "excellent", "good', "average", "poor", "very bad.") It is 1 if x is extremely high but 0 if x is extremely low. Determining the exact value of x may be quite burdensome. But determining x' will be much easier. Our task of describing information is then lightened.

Multitasking and the value of information: In Section 1.2, I first consider a multitask model in which Agent chooses a desirable action a_1 and an undesirable action a_2. To focus on the effect of information structure on contractual efficiency, I simplify that a_2 is cheap and useless so that it is zero in equilibrium. Suppose now more general case in which Agent may choose a positive level of a_2. In this case, the cosine result in Proposition 1.1 characterizes optimal contracts. An optimal contract still consists of target and bonus. Target depends directly on information structure but bonus does not. There are two complications. First, target depends on not just the sign of the likelihood ratio with respect to a_1, that is, $\frac{f_1(x|a_1,a_2)}{f(x|a_1,a_2)}$, but also of the likelihood ratio with respect to a_2, that is, $\frac{f_2(x|a_1,a_2)}{f(x|a_1,a_2)}$. An additional complication arises because target is involved with the shadow prices μ_1 and μ_2. It is therefore straightforward that "non-sign-sufficiency" with respect to a_1 is in general neither necessary nor sufficient for evaluation refinement. However, if there is a large stock of existing information so that an existing signal is a continuous variable and if signals are conditionally independent, then "non-sign-sufficiency" with respect to a_1 is sufficient for evaluation refinement, which provides Agent a grater incentive to choose a_1 and a greater disincentive to choose a_2.

Informational folk theorem: If the existing signal x is fine, then Principal can keep improving contractual efficiency by repeatedly adding sign-changing signals. Then, how much can she improve efficiency? Can she attain the first-best outcome? A positive answer is given in the following.

Proposition 1.4. *Suppose $x_i \sim_{iid} f(\cdot|a^*)$ for $i \in \mathbb{N}$. Suppose further that (i) $\frac{f_a(x_i|a^*)}{f(x_i|a^*)}$ is a continuous random variable, and (ii) the first-order approach holds for each $x^n \equiv (x_1, ..., x_n)$. Then, for any $\bar{s} \geq 0$, there exists $N \in \mathbb{N}$ such that the first-best outcome is obtained if $n \geq N$.*

Proof: Fix $a = a^*$. Suppose $s(a) \leq \bar{s}$. Then, the first-best action is zero and is trivially induced by a "null" contract. Suppose instead $s(a) > \bar{s}$. Thus, $V(a) > 0$. Let $h^n(x^n|a)$ be the distribution of iid signals $x^n = (x_1, ..., x_n)$, with h^n_a being its partial derivative with respect to a. Then, I have

$$\frac{h^n_a(x^n|a)}{h^n(x^n|a)} = \sum_{i=1}^{n} \frac{f_a(x_i|a)}{f(x_i|a)}$$

Let $b(x^n)$ be the multi-target single-bonus contract. Then,

$$Cov\left(b(x^n), \frac{h^n_a(x^n|a)}{h^n(x^n|a)}\bigg|a\right) = Cov\left(V(a)1_{\left\{\sum_{j=1}^{n}\frac{f_a(x_j|a)}{f(x_j|a)}>0\right\}}, \sum_{i=1}^{n}\frac{f_a(x_i|a)}{f(x_i|a)}\bigg|a\right)$$

$$= \sum_{i=1}^{n} V(a)\, Cov\left(\mathit{1}_{\left\{\sum_{j=1}^{n} \frac{f_a(x_j|a)}{f(x_j|a)}>0\right\}}, \frac{f_a(x_i|a)}{f(x_i|a)} \middle| a \right)$$

Since x_1, ..., x_n are identical in distribution, I have

$$= n\, V(a)\, Cov\left(\mathit{1}_{\left\{\sum_{j=1}^{n} \frac{f_a(x_j|a)}{f(x_j|a)}>0\right\}}, \frac{f_a(x_1|a)}{f(x_1|a)} \middle| a \right)$$

$$= n\, V(a) \int_{\left\{\frac{f_a(x_1|a)}{f(x_1|a)}>\kappa(x_2,...,x_n)\equiv -\sum_{j=2}^{n} \frac{f_a(x_j|a)}{f(x_j|a)}\right\}} \frac{f_a(x_1|a)}{f(x_1|a)} \prod_{j=1}^{n} f(x_j|a) dx_j$$

$$= n\, V(a) \int_{\{\kappa(x_2,...,x_n)\in(-\infty,\infty)\}} \left(\int_{\left\{\frac{f_a(x_1|a)}{f(x_1|a)}>\kappa(x_2,...,x_n)\right\}} \frac{f_a(x_1|a)}{f(x_1|a)} f(x_1|a) dx_1 \right) \prod_{j=2}^{n} f(x_j|a) dx_j$$

Since $\frac{f_a(x_j|a)}{f(x_j|a)}$ has zero mean, the inner integral above is nonnegative for any $\kappa(x_2, ..., x_n) > -\infty$, and hence, for any $\varepsilon > 0$, it follows that

$$\geq n\, V(a) \int_{\{\kappa(x_2,...,x_n)\in(0,\varepsilon)\}} \left(\int_{\left\{\frac{f_a(x_1|a)}{f(x_1|a)}>\kappa(x_2,...,x_n)\right\}} \frac{f_a(x_1|a)}{f(x_1|a)} f(x_1|a) dx_1 \right) \prod_{j=2}^{n} f(x_j|a) dx_j$$

The last line tends to infinity as n tends to infinity since the outer integral is positive: by (i), there exists $\varepsilon > 0$ such that $\{\kappa(x_2, ..., x_n) \in (0, \varepsilon)\}$ and $\left\{\frac{f_a(x_1|a)}{f(x_1|a)} > \kappa\right\}$, for each $\kappa = \kappa(x_2, ..., x_n) \in (0, \varepsilon)$, are sets of positive measure.

Therefore, the multi-target single-bonus contract can induce an arbitrarily large action. If I reduce the size of bonus appropriately, the first-best action can be induced. ■

Proposition 1.4 can be regarded as an *informational folk theorem* given $\delta \ll 1$. The virtue of Proposition 1.4 is that it is based on relational contracting so that a huge number of signals are applicable. If they are informative enough in the sense of iid condition, then Principal can and should use them to achieve the first-best outcome. Although similar informational folk theorems could be shown in formal contract models as well, their implications are limited since the set of available signals is small in formal contracting. The vice is the restrictive iid assumption. However, the iid assumption is much more sufficient than necessary for the informational folk theorem.

1.3.2 Comparison with Other Contractual Frameworks

The relational contract model is similar to formal contract models with limited liabilities and career concerns models. However, it turns out that NSS is neither necessary nor sufficient in these two frameworks.

The Limited Liability Model of Innes (1990): The model is the same as the standard formal contract model except that Agent is risk-neutral and both players' liabilities are limited as such

$$0 \leq w + b(q, x) \leq q \quad \forall q \tag{1.18}$$

where $q \in [0, \infty)$ is the output and x is a vector signal excluding q. (Note that limited liability is effective only if the output q is verifiable). Although this limited liability condition resembles the self-enforcement condition (SE), there is an important difference: self-enforcement imposes no restriction on the fixed pay w, whereas limited liability does. Agent's participation constraint then affects contractual efficiency in the limited liability model but not in the relational contract model.

Suppose that MLRP holds and that there exists an optimal contract that induces a positive action $a^f < a^*$. Let λ^f and μ^f be the Lagrange multipliers of Agent's participation and incentive constraints in equilibrium, respectively. Then, an optimal contract equals

$$w^f + b^f(q,x) = q \cdot 1_{\left\{\frac{f_a(q,x|a^f)}{f(q,x|a^f)} \geq \kappa^f\right\}}, \quad \kappa^f = \frac{1-\lambda^f}{\mu^f} > 0 \tag{1.19}$$

Since $a^f < a^*$, MLRP implies that the cutoff κ^f is positive.[11] Suppose that a new signal has the sign change effect, and hence, it also has a level change effect in some levels of the likelihood ratio. But that does not necessarily mean that it has a level change effect in a neighborhood of κ^f. In this case, the new signal is not valuable. Suppose instead that a new signal does not have the sign change effect. But it is still possible that it changes the level of the likelihood ratio near κ^f so that the new signal is valuable. Thus, "non-sign-sufficiency" is neither necessary nor sufficient for strict Pareto improvement. See Table 1.1

The Career Concerns Model of Dewatripont et al. (1999): This two-period career concerns model is similar to the relational contract model in that formal contracting is impossible, the relationship repeats and Agent is risk-neutral. Differences are that informal contracting is impossible and Principal cannot observe Agent's productivity θ. Although there is no formal/informal incentive scheme, Agent faces a work incentive: his action today contributes to his performance today, based on which Principal updates her expectation of Agent's marginal product, which is equal to the fixed wage tomorrow. Then,

[11] Note that $a^f < a^*$ means the equilibrium incentive is less than the joint-surplus maximizing incentive, where both are evaluated at $a = a^f$. That is,

$$\int_{\left\{\frac{f_a}{f} < \kappa^f\right\}} q f_a(x,q|a^f) d(q,x) = \int_{\left\{-\infty \leq \frac{f_a}{f} \leq \infty\right\}} q f_a(x,q|a^f) d(q,x) - \int_{\left\{\frac{f_a}{f} \geq \kappa^f\right\}} q f_a(x,q|a^f) d(q,x) > 0$$

Suppose $\kappa^f \leq 0$. Then, MLRP implies that $\int_{\{f_a < f\kappa^f\}} q f_a(x,q|a^f) d(q,x) \leq 0$, a contradiction.

Agent's incentive equals, for a given equilibrium action $a = a^h$,

$$Cov\left(\mathbb{E}[\theta|x, y, a^h], \frac{h_a(x, y|a)}{h(x, y|a)} \,\middle|\, a\right) \tag{1.20}$$

If y and θ are independent conditional on (x, a), then the addition of y does not affect this incentive even if it changes the sign of the likelihood ratio. If y and θ are not independent conditional on (x, a) and the addition of y changes the level of the likelihood ratio, then this incentive may increase even if y has no sign change effect. Therefore, "non-sign-sufficiency" is neither necessary nor sufficient.

Table 1.1: Information ranking with inclusive signals

Model	Trade-off	Level-change	Sign-change
Relational contracts with public signals	self-enforcement and incentive	necessary	necessary and sufficient
Agent is risk-averse (Formal contract)	risk-sharing and incentive	necessary and sufficient	sufficient
Limited liability (Formal contract)	limited liability and incentive	necessary	not necessary nor sufficient
Career concerns (No contract)		necessary	not necessary nor sufficient

1.4 Information Ranking 2: General (Public) Signals

Consider an arbitrary multi-dimensional signal $z \in Z \subset \mathbb{R}^m$, $m \in \mathbb{N}$. z may or may not be in an inclusive relation with the existing signal x. Thus, the inclusive signal case nests on this general signal case. Let $g(z|a)$ denote the probability density function of z conditional on a, where $g(z|a)$ is positive everywhere and having the partial derivative g_a.

One might conjecture that the information structure (Z, g) is more valuable (in the sense of weak Pareto improvement) than (X, f) if the former is more informative than the latter in terms of Blackwell's notion of sufficiency. Blackwell's sufficiency is most appropriate in the single-person Bayesian decision-making setting in which there are no strategic interactions. However, strategic interactions such as incentivization and self-enforcement are the keys of relational contracts. In the following, I consider a different notion of informativeness that is more general and more appropriate in the relational contract model than Blackwell's sufficiency.

Definition 1.3. (Z, g) is mean-preserving spread in likelihood ratios (or simply MPS) more informative than (X, f) if for any white noise ε such that $\mathbb{E}[\varepsilon|x, a] = 0\ \forall(x, a)$, $\frac{g_a(z|a)}{g(z|a)} = \frac{f_a(x|a)}{f(x|a)} + \varepsilon$. Equivalently, (Z, g) is MPS more informative than (X, f) if $\frac{f_a(x|a)}{f(x|a)} = \mathbb{E}\left[\frac{g_a(z|a)}{g(z|a)} \mid x, a\right]\ \forall(x, a)$.

The MPS relation above means that (Z, g) has an additional variation in likelihood ratio than (X, f) has. Both the MPS relation and Blackwell's sufficiency induce partial orders of information structures. If (Z, g) is Blackwell sufficient for (X, f), it is MPS more informative than (X, f), but not vice versa.[12] Thus, the MPS relation is more general than Blackwell's relation in the sense that the former induces a less partial order than the latter.

Note that the MPS relation is a weak order. To provide a strict order, consider the following related conditions:

$$\left\{(x, z) : \frac{f_a}{f} \neq \frac{g_a}{g}\right\} \textit{ is of positive measure} \tag{1.21}$$

[12]If (Z, g) is MPS more informative than (X, f), it is also Fisher more informative, but not vice versa.

$$\left\{(x,z):\frac{f_a}{f}>0,\frac{g_a}{g}<0\right\},\ \left\{(x,z):\frac{f_a}{f}<0,\frac{g_a}{g}>0\right\}\ \textit{or}\ \left\{(x,z):\frac{f_a}{f}=0,\frac{g_a}{g}\neq 0\right\}$$
$$\textit{is of positive measure} \tag{1.22}$$

Given that the MPS relation holds, (1.21) and (1.22) are generalizations of "non-sufficiency" and "non-sign-sufficiency," respectively. Given that the MPS relation holds, (1.22) implies (1.21), but not vice versa. Kim (1995) shows in the standard formal contract model that (Z, g) is strictly more efficient than (X, f) in terms of Pareto efficiency if the MPS relation and (1.21) are both satisfied. In the relational contract model, I need a stronger condition (1.22) to obtain such strict Pareto improvement.

Proposition 1.5. *Suppose that (P2) has solutions* $\{a^f, b^f(x)\}$ *and* $\{a^g, b^g(z)\}$ *under* (X, f) *and* (Z, g)*, respectively, and that FOA holds under both* (X, f) *and* (Z, g)*. Then, (i)* (Z, g) *weakly Pareto improves on* (X, f) *if the former is MPS in likelihood ratio more informative than the latter. (ii)* (Z, g) *strictly Pareto improves on* (X, f) *if the former is MPS in likelihood ratio more informative than the latter, (1.22) holds and* $0 < a^f < a^*$.

Proof: In the inclusive signal case, the MPS relation is always satisfied and is equivalent to the martingale property (1.14) in the proof of Proposition 1.2. Given the martingale property, "non-sign-sufficiency" and "strict Pareto improvement" are equivalent. Similarly, in this general signal case, the MPS relation and "non-sign-sufficiency" imply "strict Pareto improvement" (given that $a^f \in (0, a^*)$ so that there is a room for improvement). That is, the sufficiency part of the proof of Proposition (1.2) applies here. The converse, however, is not true in general since two general signals may not be in the MPS relation. ■

Since Principal is risk-neutral, it seems natural that she does not need to care for variation in likelihood ratios. However, Proposition 1.5 shows that actually she must care for the variation since the greater the variation, the easier it is for Principal to incentivize Agent. If I replace (X, f) with (Z, g), then Agent's work incentive $Cov\left(b^i(x), \frac{i_a}{i}\right)$ for $i = f, g$ increases (weakly) as the variation in likelihood ratio increases in the MPS sense. Given that (Z, g) can induce a (weakly) greater incentive than (X, f), the optimal target is refined if the sign-change condition (1.22) holds. Thus, by the same logic I use in the inclusive signal case, better information strictly improves efficiency.

1.5 Information Ranking 3: General Private Signals

In previous sections, I assume that signals are public. In some cases, however, it would be more reasonable to assume that only Principal can observe signals since Principal often has better access to information than Agent. Two of many cases in which signals are private are as follows. (i) Performance measures are so subjective that Agent often disagrees with Principal's evaluations. (ii) Principal uses various measures together, and thus, the weight or importance of each measure relative to others is unclear to Agent.

If the multi-signal $x = (x_1, ..., x_n)$ is Principal's private information, then she may not report the true value of x in order to save incentive payments. To give Principal the truth-telling incentive, her continuation payoff after observing x must be constant in x. To give Agent the work incentive, his continuation payoff after observing x must vary with x. Therefore, the sum of Principal's and Agent's continuation payoff must be contingent on x, which means (a sum of discounted) surplus(es) must be burned in some states of nature. See MacLeod (2003), Levin (2003) and Fuchs (2007). The necessity of surplus-burning implies that first-best outcomes cannot be achieved in general. However, Fuchs (2007) shows in a stylized model in which x and a are both binary variables that first-best outcomes can be achieved asymptotically ($\delta \to 1$) if Principal burns an unlimited amount of surplus.[13] However, burning more than the value of relationship is unreasonable.[14] Furthermore, it is also non-credible: if they are to burn such a large amount, they instead finish the relationship.

The relational contract model with the private signal x is an example of the repeated game with private monitoring in which there is no tractable recursive structure in general.[15] To make the problem tractable, Levin (2003) confines his focus to *every-period review contracts*, in which Principal reviews Agent's performance every period, and hence information is revealed in every period. He shows that every-period review contracts are not efficient for any $\delta \leq 1$. Instead, he proposes *every-T-period review contracts*, in which Principal evaluates Agent every T periods based on performance over the last T periods, which means that Principal hides information for $T-1$ periods. He conjectures that it is asymptotically efficient as $\delta \to 1$. Fuchs (2007) confirms this conjecture in his binary signal and binary action model. However, every-T-period review contracts are still suboptimal in general for $\delta < 1$. In his stylized model, Fuchs (2007) shows that an efficiency wage contract is optimal:

[13] Let $x_t \in \{0,1\}$ and $a_t \in \{0,1\}$ for each t. Consider the following contract. Principal reviews Agent's performances $(x_1, ..., x_T)$ up to some fixed date T. Before T, there is no evaluation. If $x_t = 0$ for all $t \leq T$, then Principal burns some money, but not otherwise. If $\delta \to 1$ and Principal lets T tend to ∞, then burning an infinite amount of surplus in the infinite future results in approximating the first-best outcome arbitrarily closely.

[14] Of course, if $\delta \to 1$, then the value of relationship tends to infinity, and therefore, burning an unlimited amount of surplus may not necessarily imply burning more than the value of the relationship. However, in this paragraph, I consider any $\delta \in (0,1)$.

[15] See Kandori (2002).

Agent works (rather than shirks) and is paid a fixed wage every period until he is fired. Principal gives no feedback (about Agent's performance) to Agent at all. This no-feedback property is quite inconsistent with real-life business relationships, in which employers regularly evaluate employees' performances.

In this section, I extend the model of Fuchs (2007) to continuous action and continuous multi-signal. An important advantage with continuous signal is that contracts could have non-degenerate rich structures. Thus, I can learn how the privateness of the signal affects optimal contracting. Considering this continuous/multi-dimensional model is also advantageous since it is a building block upon which I can develop an information theory.

I confine our focus to every-T-period review contracts because of their tractability, asymptotic efficiency and consistency with real-life business relationships. For a given T, Principal makes the following promise. At the end of each period $t = 1, ..., T-1$, Principal pays w_t to Agent. At the end of period T, upon observing $(x_1, ..., x_T)$, she pays $w_T - \lambda(x_1, ..., x_T)$ to Agent and burns the remaining $\lambda(x_1, ..., x_T)$. Thus, Principal's spending in period T is independent of $(x_1, ..., x_T)$. Given this promise, Agent's expected utility in each period $t = 1, ..., T$ equals $w_t - \delta^{T-t}\mathbb{E}[\lambda(x_1, ..., x_T)|a_1, ..., a_T] - c(a_t)$. Then, his work incentive equals $\frac{\partial}{\partial a_t}\delta^{T-t}\mathbb{E}[-\lambda(\cdot, x_t, \cdot)|\cdot, a_t, \cdot]$. Since Principal has full bargaining power, she will choose w_t so that Agent's payoff equals $\bar{u}^A$ in every period:

$$w_t = \delta^{T-t}\mathbb{E}[\lambda(x_1, ..., x_T)|a_1, ..., a_T] + c(a_t) + \bar{u}^A \qquad \text{(BIR2)}$$

Principal's payoff for the T-period then equals $\sum_{t=1}^{T}\delta^{t-1}\{\mathbb{E}[q_t|a_t] - w_t\} = \sum_{t=1}^{T}\delta^{t-1}\{\mathbb{E}[q_t|a_t] - c(a_t)\} - \delta^{T-1}\mathbb{E}[\lambda(x_1, ..., x_T)|a_1, ..., a_T]$ up to a constant. Note that it is the total surplus for the T-period. Let s_T denote this total surplus. Since Principal offers such a promise every T periods, her payoff equals

$\frac{1}{1-\delta^T}s_T$ and the value of (ongoing) relationship is $\frac{\delta}{1-\delta^T}\left[s_T - \sum_{t=1}^{T}\delta^{t-1}\bar{s}\right]$. Note that Principal will keep her promise of burning the surplus if $\lambda(x_1, ..., x_T)$ is no greater than the value of relationship. Let f_t denote f's partial derivative in a_t. Therefore, for each given T, Principal solves the following problem:

$$\max_{(a_1,...,a_T),\lambda(x_1,...,x_T)} \frac{s_T}{1-\delta^T} = \frac{1}{1-\delta^T}\left[\sum_{t=1}^{T}\delta^{t-1}\{\mathbb{E}[q_t|a_t] - c(a_t)\} - \delta^{T-1}\mathbb{E}[\lambda(x_1, ..., x_T)|a_1, ..., a_T]\right] \tag{P3}$$

s.t.

$$\delta^{T-t}Cov\left(-\lambda(\cdot, x_t, \cdot), \frac{f_t(x_t|a_t)}{f(x_t|a_t)}\,\middle|\, \cdot, a_t, \cdot\right) = c'(a_t), \ \ t \in 1, ..., T \tag{IC2}$$

$$0 \le \lambda(x_1, ..., x_T) \le \frac{\delta}{1-\delta^T}\left[s_T - \sum_{t=1}^{T}\delta^{t-1}\bar{s}\right] \equiv V_T \quad \forall (x_1, ..., x_T) \tag{SE2}$$

Then, Principal chooses $T = T^f$ such that $T^f \in \arg\sup_{T\in\{1,...,\infty\}} \frac{s_T}{1-\delta^T}$. Sometimes, business norms constrain the set of feasible T. Principal then chooses a constrained optimum T^f.

Proposition 1.6. *Suppose (i) the first-order approach is valid, (ii) there exists an optimal every-T-period review contract that solves (P3) such that the solution induces an action vector* $(a_1^f, ..., a_T^f) \gg (0, ..., 0)$ *with* T^f *and* $\frac{1}{1-\delta^{T^f}}s_{T^f}^f$ *being the optima of* T *and* $\frac{s_T}{1-\delta^T}$*, respectively. Then, an optimal surplus-burning contract is given by*

$$\lambda_T^f(x) = V_T^f \cdot 1_{\left\{\sum_{t=1}^{T}\left\{\mu_t^f \delta^{T-t}\frac{f_t(x_t|a_t^f)}{f(x_t|a_t^f)}\right\} < -\psi_T^f\right\}} \quad at \ \ T = T^f \tag{1.23}$$

where $V_T^f \equiv \frac{\delta}{1-\delta^T}\left[s_T^f - \sum_{t=1}^{T}\delta^{t-1}\bar{s}\right] > 0$, $\psi_{T^f}^f > 0$, $\mu_t^f > 0 \ \forall t$, $\sum_{t=1}^{T^f}\mu_t^f\delta^{T^f-t} = 1$ *and* $(a_1^f, ..., a_{T^f}^f)$ *is given by (IC2).*

Proof: For notational simplicity, let $x = (x_1, ..., x_T)$. Let μ_t, $l(x)$ and $g(x)$ be the Lagrange multipliers for (IC2) and the two inequalities in (SE2), respectively. Then, the Kuhn-Tucker condition with respect to $\lambda(x)$ is given by

$$\prod_{k=1}^{T} f(x_k|a_k) \cdot \left[\frac{\delta^{T-1}}{1-\delta^T}\left(1+\delta\int g(u)du\right)+\sum_{t=1}^{T}\mu_t\delta^{T-t}\frac{f_t(x_t|a_t)}{f(x_t|a_t)}\right] = l(x) - g(x) \tag{1.24}$$

Thus, $\lambda(x)$ equals 0 if the sum of terms in bracket is positive. $\lambda(x)$ equals V_T if the sum is negative. If the sum is zero, then $\lambda(x)$ can be any level in between $[0, V_T]$ since that $(a_1, ..., a_T) \gg (0, ..., 0)$ implies $V_T > 0$. If I let $\psi_T = \frac{\delta^{T-1}}{1-\delta^T}\left(1+\delta\int g(u)du\right)$, then (1.23) is proven. Clearly, $\psi_T > 0$.

Suppose $\mu_t \leq 0$ for some $t = k \in \{1, ..., T\}$. $\lambda(\cdot, x_k, \cdot)$ is nondecreasing in $\frac{f_k(x_k|a_k)}{f(x_k|a_k)}$ by (1.23). The covariance term in (IC2) at $t = k$ is nonpositive, which contradicts $a_k > 0$.

Since $1_{\{\sum_{t=1}^{T}\mu_t M_t < -N\}} = 1_{\{\sum_{t=1}^{T}\alpha\mu_t M_t < -\alpha N\}}$ for any $M_t, N, \alpha \in (0, \infty)$, we can scale $(\{\mu_t\}, \psi_T)$ so that $\sum_{t=1}^{T}\mu_t^f\delta^{T-t} = 1$. ■

Hence, it is optimal to burn the whole value of the relationship $V_{T^f}^f$ in the last period T^f if and only if a time-weighted average of likelihood ratios $\sum_{t=1}^{T^f}\left\{\mu_t^f\delta^{T^f-t}\frac{f_t(x_t|a_t^f)}{f(x_t|a_t^f)}\right\}$ is less than a negative cutoff level $-\psi_{T^f}^f$. Such a dichotomous burning contract can be implemented by a termination contract with an efficiency wage: in each period $t = 1, ..., T^f$, Agent is paid an efficiency wage w_t^f such that

$$w_t^f = \delta^{T^f-t}\mathbb{E}[\lambda^f(x_1, ..., x_{T^f})|a_1^f, ..., a_{T^f}^f] + c(a_t^f) + \bar{u}^A \tag{1.25}$$

Note that w_t^f is greater than the reservation wage $\bar{u}^A$. In period T^f, if Agent passes the review, then Principal continues the relationship and pays him $w^f_{T^f}$. If he fails, then Principal permanently terminates the relationship by firing him (without rehiring) and pays him only $w^f_{T^f} - V^f_{T^f}$. Permanent termination means the whole value of relationship disappears. In the following, I interpret the optimal burning contract as the termination contract.

As time goes by (from 1 to T^f), Agent faces greater incentives, as the punishment at T^f becomes more serious due to discounting. That is, Principal faces less difficulty in incentivization as times passes. However, the value of action a_t, evaluated at the initial period, for Principal is decreasing in t due to discounting. To balance these countervailing effects, Principal puts appropriate weights $\{\mu_t\}$ on the discounted likelihood ratio $\{\delta^{T-t}\frac{f_t}{f}\}$.

The optimal review period T^f is finite with $\delta < 1$. Suppose not. As $T \to \infty$, (IC2), (1.23) and $a^f_1 > 0$ imply that $V_T \to \infty$, which contradicts (SE2). If T^f is finite, then it means that the welfare loss from termination is unavoidable. In contrast, as $\delta \to 1$, the value of relationship becomes unlimited, and hence, Agent's work incentive in (IC2) becomes unbounded. Principal can thus make a contract, which is less powerful than $\lambda^f_T(x)$, that induces the first-best action in every period. Then, by letting T tend to ∞, Principal's payoff approaches the first-best discounted sum of surpluses since the expected loss from termination in period T is negligible.

Information Structure Comparison

In the general case with $\delta < 1$, the welfare loss from termination is inevitable. To reduce the termination probability, Principal will use more generous evaluation criteria than she uses in the public signal case: the cutoff of the likelihood ratio is zero with public signals, whereas it is a negative value $-\psi^f_{T^f}$ with private signals. Since evaluation is more generous, Agent faces a weaker incentive. That is, Principal must endure some loss in productivity (i.e., lower action) in order to make the relationship more sustainable.

However, this conflict between productivity and sustainability can be alleviated, for a given $\delta \ll 1$, if Principal can choose information structure. The following proposition shows that one information structure is more valuable than another if the former is MPS more informative than the latter. The proof constructively demonstrates that the former is more efficient in dealing with the conflict. The intuition and idea of the proof are as follows. The optimal contract $\lambda^i_{T^i}(x)$, $i = f, g$, consists of three components: the review criteria, the size of punishment, and the review period. The information structure directly affects the review criteria, but only indirectly affects the other two components. As Principal changes the information structure from $\frac{f_t}{f}$ to $\frac{g_t}{g}$, the variation in likelihood ratio increases. This increased variation enables Principal to refine the review criteria in such a way that the work incentive is unaltered even if evaluation is more generous. As the trade-off between the productivity and sustainability of the relationship is relaxed, Principal may extend review period T so that the relationship lasts longer.

Proposition 1.7. *Suppose the first-order approach is valid and (P3) has solutions under* (X, f) *and* (Z, g)*, respectively, where* (X, f) *induces* $(a^f_1, ..., a^f_{T^f}) \gg (0, ..., 0)$*. Then, (i)* (Z, g) *weakly Pareto improves on* (X, f) *if the former is MPS in likelihood ratio more informative than the latter. (ii) In addition to the assumptions above, if* $T^f = 1$ *and* $\int^l R^g_a(u)du > \int^l R^f_a(u)du \ \forall l < 0$*, where* $R^i_a(l) \equiv \mathbb{P}(\frac{i_a}{i} \leq l|a)$ *is the distribution of* $\frac{i_1}{i} = \frac{f_1}{f}, \frac{g_1}{g}$*, then strict Pareto improvement is obtained.*

Proof: I fix $T = T^f$ and shall show that (Z, g) improves on (X, f) even if Principal chooses $T^g = T^f$. Then, I suppress the subscript T because it does not play any role in the proof. I can rewrite the bonus amount $\lambda^i(\cdot)$ for $i = g, f$ as such $\lambda^i(\cdot) = V_T 1_{\left\{1+\sum_{t=1}^T \left\{\nu_t^i \delta^{T-t} \frac{i_t(\cdot|a)}{i(\cdot|a)}\right\}<0\right\}}$ where $\nu_t^i \equiv \frac{\mu_t^i}{\psi^i}$.

The following new but related problem is useful in this proof:

$$\min_{\lambda(x_1,...,x_T)} \mathbb{E}^i[\lambda(x_1, ..., x_T)|a] \textit{ s.t. (IC2) and (SE2) for given } (a_1, ..., a_T) > 0 \textit{ and } V_T \tag{P4}$$

where E^i is expectation with respect to information structure $i = f, g$. It can be easily shown that for a constant $\hat{\nu}_t^i > 0$, a solution $\hat{\lambda}^i(\cdot)$ to (P4) is

$$\hat{\lambda}^i(\cdot) = V_T 1_{\left\{1+\sum_{t-1}^T \left\{\hat{\nu}_t^i \delta^{T-t} \frac{i_t(\cdot|a)}{i(\cdot|a)}\right\}<0\right\}} \tag{1.26}$$

If I solve (P4) for given $(a_1^i, ..., a_T^i) \gg 0$ and V_T^i, then $(a_1^i, ..., a_T^i, \hat{\lambda}^i(\cdot))$ solves (P3). Suppose

$$\mathbb{E}^g[\hat{\lambda}^g(z)|a] \leq \mathbb{E}^f[\hat{\lambda}^f(x)|a] \textit{ for any given } a = (a_1, ..., a_T), V_T \tag{1.27}$$

Then, $s_T^g \geq s_T^f$. Thus, it is sufficient to show weak inequality in (1.27) for (i) and strict inequality for (ii). For notational convenience, I write $\mathbb{E}^i[\cdot]$ for $\mathbb{E}^i[\cdot|a_1, ..., a_T]$.

Let $r(q^i) \equiv V_T 1_{\{q^i<0\}}$, $q^i \equiv \left(1 + \sum_{t=1}^T \left\{\hat{\nu}_t^i \delta^{T-t} \frac{i_t(\cdot|a)}{i(\cdot|a)}\right\}\right)$. (IC2) and (1.26) imply that $c'(a_t) = -\delta^{T-t}\mathbb{E}^i\left[\hat{\lambda}^i(\cdot)\frac{i_t(\cdot|a)}{i(\cdot|a)}\right] = -\delta^{T-t}\mathbb{E}^i\left[r(q^i)\frac{i_t(\cdot|a)}{i(\cdot|a)}\right] \quad \forall i = g, f$ since each a_t is fixed. Thus,

$$\mathbb{E}^i[r(q^i)q^i] = \mathbb{E}^i\left[r(q^i)\left(1 + \sum_{t=1}^T \left\{\hat{\nu}_t^i \delta^{T-t} \frac{i_t(\cdot|a)}{i(\cdot|a)}\right\}\right)\right] = \mathbb{E}^i[r(q^i)] - \sum_{t=1}^T \left\{\hat{\nu}_t^i c'(a_t)\right\} \tag{1.28}$$

Let $\zeta(q^i) \equiv r(q^i)q^i$. Then, $\zeta(q^i) = 0 \ \forall q^i \geq 0$ and $\zeta(q^i) = V_T \cdot q^i$ $\forall q^i < 0$. That is, $\zeta(q^i)$ is a (nonlinear) concave function. $\zeta(q^i)$ is non-differentiable only at $q^i = 0$. However, without loss of generality, I can assume that $\zeta'(0) = V_T$. Then, $\zeta'(q^i) = r'(q^i)q^i + r(q^i) = r(q^i)$. Note also that $\zeta\left(1 + \sum_{t=1}^{T}\left\{\hat{\nu}_t \delta^{T-t}\frac{i_t(\cdot|a)}{i(\cdot|a)}\right\}\right)$ is concave in $\hat{\nu}_t$. Let $q' \equiv \left(1 + \sum_{t=1}^{T}\left\{\hat{\nu}_t^g \delta^{T-t}\frac{f_t(x|a)}{f(x|a)}\right\}\right)$. Then, by concavity,

$$\mathbb{E}^f[\zeta(q^f)] - \mathbb{E}^f[\zeta(q')] \geq \mathbb{E}^f[\zeta'(q^f)(q^f - q')] = \mathbb{E}^f[r(q^f)(q^f - q')] = \sum_{t=1}^{T}\left\{(\hat{\nu}_t^g - \hat{\nu}_t^f)c'(a_t)\right\} \tag{1.29}$$

where the second equality is by (1.28).

(i) Let $\rho_t^i(l_t)$ be the distribution of $l_t = \frac{i_t(\cdot|a)}{i(\cdot|a)} \in \mathbb{R}$ for $i = g, f$. Then, it follows that

$$\mathbb{E}^g[\hat{\lambda}^g(y)] - \mathbb{E}^f[\hat{\lambda}^f(x)] = \mathbb{E}^g[r(q^g)] - \mathbb{E}^f[r(q^f)] = \mathbb{E}^g[\zeta(q^g)] - \mathbb{E}^f[\zeta(q^f)] + \sum_{t=1}^{T}\left\{(\hat{\nu}_t^g - \hat{\nu}_t^f)c'(a_t)\right\}$$

$$\leq \mathbb{E}^g[\zeta(q^g)] - \mathbb{E}^f[\zeta(q^f)] + \left(\mathbb{E}^f[\zeta(q^f)] - \mathbb{E}^f[\zeta(q')]\right) = \mathbb{E}^g[\zeta(q^g)] - \mathbb{E}^f[\zeta(q')]$$

$$= \int_{\mathbb{R}^T} \zeta\left(1 + \sum_{t=1}^{T}\left\{\hat{\nu}_t^g \delta^{T-t} l_t\right\}\right)\left(\prod_{t=1}^{T}\rho_t^g(l_t) - \prod_{t=1}^{T}\rho_t^f(l_t)\right)\prod_{t=1}^{T} dl_t$$

$$= \sum_{t=1}^{T}\int_{\mathbb{R}^{T-1}}\left\{\int_{\mathbb{R}}\zeta\left(1 + \hat{\nu}_t^g\delta^{T-t}l_t + \sum_{s\neq t}\left\{\hat{\nu}_s^g\delta^{T-s}l_s\right\}\right)\left(\rho_t^g(l_t) - \rho_t^f(l_t)\right)dl_t\right\} \times \prod_{k<t}\rho_k^g(l_k)\prod_{k>t}\rho_k^f(l_k)\prod_{k\neq t}dl_k \leq 0$$

where $\int_{\mathbb{R}^T}$ denotes the integral on $\mathbb{R}^T$. The first equality is by definition of $r(\cdot)$ and (1.26), the second equality is by (1.28), and the first inequality is by

(1.29). For the last equality, see the footnote.[16] The integral in brackets $\{\cdot\}$ is nonpositive for all t since ρ_t^g is a mean-preserving spread of ρ_t^f and $\zeta(\cdot)$ is a concave function of l_t.[17] Then, the last inequality holds, and hence, I get the desired result (1.27).

(ii) Let $T = 1$. Then, $\mathbb{E}^g[\zeta(q^g)] - \mathbb{E}^f[\zeta(q')] = \int_L \zeta\left(1 + \hat{\nu}_1^g l_1\right)\left(\rho_1^g(l_1) - \rho_1^f(l_1)\right) dl_1$. It is equivalent to (through integration by parts) $-V_T \hat{\nu}_1^g \left(\int^{-1/\hat{\nu}_1^g} [R_a^g(u) - R_a^f(u)] du\right)$, which is strictly negative since $a > 0$ implies that $V_T > 0$. ■

Proposition 1.7 (i) implies that MPS in likelihood ratio is a valid information ranking order in the relational contract model with private signals. Proposition 1.7 (ii) provides an example in which MPS ordering is strict. Recall that MPS ordering is also valid in the standard formal contract model and the relational contract model with public signals. It can also be shown that this ordering is still valid in the limited liability model of formal contracts.[18] Thus, various contractual frameworks are similar in their use of general sig-

[16] To illustrate how the last equality holds, let $T = 2$, for instance. Then, $\mathbb{E}^g[\zeta(q^g)] - \mathbb{E}^f[\zeta(q')]$ equals

$$\int_{L^2} \zeta\left(1 + \hat{\nu}_1^g \delta l_1 + \hat{\nu}_2^g l_2\right)\left(\rho_1^g(l_1)\rho_2^g(l_2) - \rho_1^f(l_1)\rho_2^f(l_2)\right) dl_1 dl_2$$

$$= \int_{L^2} \zeta\left(1 + \hat{\nu}_1^g \delta l_1 + \hat{\nu}_2^g l_2\right)\left(\rho_1^g(l_1)\rho_2^g(l_2) - \rho_1^f(l_1)\rho_2^f(l_2) + \rho_1^g(l_1)\rho_2^f(l_2) - \rho_1^g(l_1)\rho_2^f(l_2)\right) dl_1 dl_2$$

$$= \int_L \left\{\int_L \zeta\left(1 + \hat{\nu}_1^g \delta l_1 + \hat{\nu}_2^g l_2\right)\left(\rho_1^g(l_1) - \rho_1^f(l_1)\right) dl_1\right\} \rho_2^f(l_2) dl_2$$

$$+ \int_L \left\{\int_L \zeta\left(1 + \hat{\nu}_1^g \delta l_1 + \hat{\nu}_2^g l_2\right)\left(\rho_2^g(l_2) - \rho_2^f(l_2)\right) dl_2\right\} \rho_1^g(l_1) dl_1$$

[17] See Rothschild and Stiglitz (1970).

[18] See (1.19). The optimal formal contract with limited liability and the optimal relational contract with private signals have the same form if the review period is $T = 1$. Thus, one can show the desired result by using the technique I employ in the proof for Proposition 1.7.

nals. This is no longer true for the career concerns model since there is no contract, only a learning mechanism. See Table 1.2.

Table 1.2: Information ranking with general signals

Model	Trade-off	MPS in Score
Relational contracts with private signals	truthful evaluation and incentive	sufficient
Relational contracts with public signals	self-enforcement and incentive	sufficient
Agent is risk-averse (Formal contract)	risk-sharing and incentive	sufficient
Limited liability (Formal contract)	limited liability and incentive	sufficient
Career concerns (No contract)		not sufficient in general

Information also affects the efficiency wage. Recall from (1.25) that $\{w_t^i\}_{t=1}^{T-1}$ are efficiency wages for $i = f, g$. Given an action a^g, the wage compression $w_t^i - \bar{u}^A$, that is, the gap between the wage inside the firm and the wage outside the firm, is reduced as the information quality is improved.

If I assume that the longer relationships continue the better information the parties acquire, then follows an interesting relationship dynamics: it is increasingly difficult to break relationships as time goes by. Consider some undesirable relationships (such as collusion, workplace politics, or organized crime). In these relationships, players cannot rely on court-enforcing formal contracts. Thus, relational contracts are natural alternatives. Information is good for them but bad for society since information improves the sustainability of these bad relationships. Therefore, early government interventions are requested to prevent these relationships from growing too strong to be broken.

1.6 First-Order Approach

Mirrlees (1976, 1999) points out that justifying the first-order approach (FOA) is neither easy nor insignificant. Since his seminal work, a large body of literature arises.[19] Rogerson (1985) shows that the MLRP-CDFC condition (see Assumptions 1.1 and 1.2), which is now standard in the literature, validates the use of FOA in a single-signal formal contract model.[20] Jewitt (1988) finds a different sufficient condition based on a formal contract model in which there are *two independent signals*. Recently, Conlon (2009) has extended Jewitt's condition to a formal contract model in which there are *multiple signals*.

However, established conditions are built on the standard formal contract model, which is systematically different from the relational contract model. Hence, one must verify whether these conditions can also be applied to the relational contract model. Furthermore, these conditions are excessively restrictive. For instance, almost every well-known probability distribution function fails to satisfy the CDFC condition.[21] Additionally, existing conditions are increasingly restrictive as the dimension n of the vector x rises. For instance, Conlon (2009) assumes that the likelihood ratio $\frac{f_a(x_1,...,x_n|a)}{f(x_1,...,x_n|a)}$ is pointwise nondecreasing for each x_i, $i = 1, ..., n$. If I use these strong conditions to justify FOA, then the set of information structures that can be used in the model is small. Information ranking orders then have limited applicability.

I use the following two-step procedure to justify FOA. First, I characterize optimal contracts in the relaxed problem (P2) without assuming the

[19]See Mirrlees (1976), Grossman and Hart (1983), Rogerson (1985), Jewitt (1988), Sinclair-Desgagne (1994), Sinclair-Desgagne (2009), and Conlon (2009).

[20]See Assumptions 1.1 and 1.2.

[21]For instance, suppose that $x = a + \epsilon$, ϵ follows a density function $p(\cdot)$ and $c(a) = a$. Then, CDFC requires $p(\epsilon)$ to be nondecreasing in ϵ.

validity of FOA. These contracts are also optimal in the original problem (P1), that is, FOA is valid, if Agent's objective function is concave in action given these contracts. Proposition 1.8 provides a sufficient condition under which Agent's objective function is concave.

Proposition 1.1 implies that optimal contracts of (P2) are given by

$$b^f(x) = \begin{cases} 0 & if \;\; \mu \frac{f_a(x|a^f)}{f(x|a^f)} < 0 \\ \in \left[0, V(a^f)\right] & if \;\; \mu \frac{f_a(x|a^f)}{f(x|a^f)} = 0 \\ V(a^f) & if \;\; \mu \frac{f_a(x|a^f)}{f(x|a^f)} > 0 \end{cases} \tag{1.30}$$

The sign of μ is nonnegative. To see this, suppose that $\mu < 0$. Then, $Cov\left(b^f(x), \frac{f_a(x|a^f)}{f(x|a^f)} \mid a^f\right) \leq 0$ since $b^f(x)$ is nonincreasing in $\frac{f_a(x|a^f)}{f(x|a^f)}$. However, $a^f > 0$ and (FOC) imply that

$$0 = c'(0) < c'(a^f) = Cov\left(b^f(x), \frac{f_a(x|a^f)}{f(x|a^f)} \,\middle|\, a^f\right)$$

which is a contradiction. Therefore, we have $\mu \geq 0$. If $\mu = 0$, then $a^f = a^*$. If $\mu > 0$, however, one cannot determine whether a^f is higher or lower than a^* without assuming that FOA is valid. To see this, consider the following Kuhn-Tucker condition with respect to a:

$$s'(a)\left[1 + \frac{\delta}{1-\delta}\int G(x)dx\right] + \mu\left(\int b(x)f_{aa}(x|a)dx - c''(a)\right) = 0 \tag{1.31}$$

which means $s'(a^f) = 0$ if $\mu = 0$ (since the Lagrange multiplier $G(x)$ is nonnegative), and hence, $a^f = a^*$. If $\mu > 0$, then $a^f < a^*$ if the terms in parenthesis is negative (i.e. the second order condition with respect to a holds), which is the case if FOA is valid.

Proposition 1.8. *Suppose the following condition holds.*

$$\mathbb{P}\left(\frac{f_a(x|a^f)}{f(x|a^f)} \leq 0 \,\middle|\, c^{-1}(\kappa)\right) \;\; \textit{is convex in } \kappa \tag{LCDFC}$$

where $a^f > 0$ is an equilibrium action of (P2) and $\kappa = c(a)$. Then, the first-order approach is valid. That is, if there is a solution to (P2) that induces a positive action a^f, it is a solution to (P1).

Proof: Since $a = c^{-1}(\kappa)$, the problem of choosing a is equivalent to the problem of choosing κ. Let $(a^f, b^f(x))$ be a solution to (P2) such that $a^f > 0$.

Suppose $a^f \neq a^*$. Then, since '$\mu \geq 0$' and '$\mu = 0$ implies $a^f = a^*$,' it follows that $\mu > 0$. (1.30) then implies that $V(a^f)1_{\left\{\frac{f_a(x|a^f)}{f(x|a^f)}>0\right\}}$ is an optimal contract. Generically, it is the unique optimal contract. Then, I can assume without loss of generality that $b^f(x) = V(a^f)1_{\left\{\frac{f_a(x|a^f)}{f(x|a^f)}>0\right\}}$. Agent's objective function then equals (ignoring the constant w^f)

$$V(a^f)\mathbb{P}\left(\frac{f_a(x|a^f)}{f(x|a^f)} > 0 \,\middle|\, a\right) - c(a) = V(a^f)\mathbb{P}\left(\frac{f_a(x|a^f)}{f(x|a^f)} > 0 \,\middle|\, c^{-1}(\kappa)\right) - \kappa$$

which is concave in κ if LCDFC is satisfied.

Suppose instead $a^f = a^*$. In this case, $b^f(x)$ may be different from the multi-target single-bonus contract (see Proposition 1.1). Then, it follows that

$$\begin{aligned}
c'(a^*) &= Cov\left(b^f(x), \frac{f_a(x|a^*)}{f(x|a^*)} \,\middle|\, a^*\right) \\
&= \int_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}>0\right\}} b^f(x)\frac{f_a(x|a^*)}{f(x|a^*)}f(x|a^*)dx + \int_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}\leq 0\right\}} b^f(x)\frac{f_a(x|a^*)}{f(x|a^*)}f(x|a^*)dx \\
&\leq \int_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}>0\right\}} V(a^f)\frac{f_a(x|a^*)}{f(x|a^*)}f(x|a^*)dx + \int_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}\leq 0\right\}} 0\frac{f_a(x|a^*)}{f(x|a^*)}f(x|a^*)dx \\
&= V(a^*)Cov\left(1_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}>0\right\}}, \frac{f_a(x|a^*)}{f(x|a^*)} \,\middle|\, a^*\right)
\end{aligned}$$

Then, there is $\hat{V} \in (0, V(a^*)]$ such that $c'(a^*) = \hat{V}Cov\left(1_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}>0\right\}}, \frac{f_a(x|a^*)}{f(x|a^*)} \mid a^*\right)$. Thus, a new contract $\hat{V}1_{\left\{\frac{f_a(x|a^*)}{f(x|a^*)}>0\right\}}$ is feasible and induces a^* and is hence optimal. Then, Agent's objective function equals (ignoring the constant w^f)

$\hat{V}\mathbb{P}\left(\frac{f_a(x|a^*)}{f(x|a^*)} > 0 \,\middle|\, c^{-1}(\kappa)\right) - \kappa$, which is concave in κ by LCDFC. ■

To get an idea of LCDFC, suppose that x is a scalar and that MLRP holds. In this case, LCDFC is reduced to $F(x|c^{-1}(\kappa))$ is convex in κ **at a single point** $x = x^f$ where x^f is such that $\frac{f_a(x^f|a^f)}{f(x^f|a^f)} = 0$. In contrast, (Levin's version of) CDFC requires $F(x|c^{-1}(\kappa))$ is convex in κ **for all** x. That is, LCDFC is a local version of CDFC.

CDFC is notoriously strong so that is inconsistent with almost every well-known distributions such as normal and exponential family. Nevertheless, Levin (2003) assumes MLRP and CDFC in this single-signal relational contract model. In the following, I provide examples that are consistent with LCDFC but not CDFC.

Example 1.4. *Suppose that $x = a + \varepsilon$ where $\varepsilon \in \mathbb{R}$ follows a distribution $Q(\cdot)$ with a density $q(\cdot)$ such that $q(\varepsilon) > 0\ \forall\varepsilon$. Suppose also that MLRP holds. Then, we have $F(x|\, c^{-1}(\kappa)) = Q(x - c^{-1}(\kappa))$, and hence, it follows that*

$$\frac{\partial^2}{\partial\kappa^2}F(x|c^{-1}(\kappa)) = \frac{q(x-a)}{\{c'(a)\}^2}\left[\frac{c''(a)}{c'(a)} + \frac{q'(x-a)}{q(x-a)}\right] \tag{1.32}$$

where $a = c^{-1}(\kappa)$.

(i) Normal error and quadratic cost: Suppose $\varepsilon \sim N(0, \sigma^2)$ and $c(a) = \frac{\gamma}{2}a^2$. In this case, $\frac{q'(x-a)}{q(x-a)} = -\frac{x-a}{\sigma^2}$. (1.32) thus implies that CDFC holds if the term in bracket $[\cdot]$ is positive, that is, $\sigma^2 > a(x-a)$, for all x for each a. Thus, CDFC does not hold if x tends to ∞. In contrast, LCDFC holds. To see this, note that $\frac{f_a(x|a^f)}{f(x|a^f)} = \frac{x-a^f}{\sigma^2}$. Thus, $\frac{f_a(x|a)}{f(x|a)}$ changes sign as x increases from below to above of a^f. Note also that $a(a^f - a)$ is maximized at $a = \frac{a^f}{2}$. Finally, LCDFC holds if $\sigma^2 > \frac{(a^f)^2}{4}$. Note that a^f depends on parameters such as σ^2. But since a^f is a constant, CDFC still holds if σ^2 is large enough.

(2) Arbitrary error, exponential cost and a compact set of actions: Suppose $c(a) = e^{\gamma a} - 1$, $a \in [0, \bar{a}]$ *and* $\bar{a} < \infty$. *LCDFC holds if* $\gamma \geq \max_{a \in [0,\bar{a}]} \left\{ -\frac{q'(a^f - a)}{q(a^f - a)} \right\}$. *Note that this is true for any error distribution* $Q(\cdot)$. *Therefore, most of exponential family distributions satisfy LCDFC if* γ *is sufficiently large.*

(3) Multivariate normal error and quadratic cost: Suppose $x = \mu a + \varepsilon$, *where* $x, \mu, \varepsilon \in \mathbb{R}^n$ *and* $\varepsilon \sim N(0, \Sigma)$. *Note that*

$$f(x|a) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(x - \mu a)' \Sigma^{-1} (x - \mu a) \right)$$

$$\frac{f_a(x|a^f)}{f(x|a^f)} = x' \Sigma^{-1} \mu - \mu' \Sigma^{-1} \mu a^f \leq 0 \quad \Leftrightarrow \quad a + \frac{\varepsilon' \Sigma^{-1} \mu}{\mu' \Sigma^{-1} \mu} \leq a^f$$

Let $v \equiv \frac{\varepsilon' \Sigma^{-1} \mu}{\mu' \Sigma^{-1} \mu} \sim N(0, \sigma_v^2)$, *where* $\sigma_v^2 \equiv \frac{1}{\mu' \Sigma^{-1} \mu}$. *Note that* $T \equiv a + v$ *is a sufficient statistic of* x. *Thus, LCDFC holds if* $\mathbb{P}(v \leq a^f - c^{-1}(\kappa))$ *is convex in* κ. *Let* $c(a) = \frac{\gamma}{2} a^2$. *Then, the previous analysis analogously implies that LCDFC holds if* $\sigma_v^2 > \frac{(a^f)^2}{4}$.

I could obtain this weak restriction LCDFC because the functional form of optimal relational contracts is known. In contrast, in Holmstrom (1979) standard formal contract model, functional forms of optimal formal contracts are unknown in general. Due to this lack of knowledge about the shape of optimal contracts, one should impose stronger conditions on information structure to validate FOA in formal contracting.

Finally, I compare LCDFC to a FOA-justifying condition recently proposed by Conlon (2009) in multisignal formal contracting context. Suppose that x is a vector signal. His condition is called the concavity of increasing set property (CISP). If MLRP holds, LCDFC is sort of a local version of CISP. I shall see this point in the following. A set $E \subset \mathbb{R}^n$ is defined as an increasing set, if for $x, x' \in \mathbb{R}^n$ and $x' \geq x$ (in the sense of component-wise order), $x \in E$ implies $x' \in E$. He defines that a density function $f(x|a)$ satisfies CISP if $\mathbb{P}(x \in E|a)$ is concave in a **for all increasing set** E. Let $E^f \equiv \left\{x : \frac{f_a(x|a^f)}{f(x|a^f)} > 0\right\}$. Then, MLRP implies that E^f is an increasing set. LCDFC requires that $\mathbb{P}(x \in E|\, c^{-1}(\kappa))$ is concave in κ **at one increasing set** $E = E^f$. Thus, LCDFC is a sort of local version of CISP. Furthermore, LCDFC is an exact local version of CISP if $\mathbb{P}(x \in E|a)$ is nondecreasing in a since $c^{-1}(\kappa)$ is concave and concavity is closed under nondecreasing and concave transformation. Conlon (2009) defines that $f(x|a)$ satisfies the non-decreasing increasing set property (NISP) condition if $\mathbb{P}(x \in E|a)$ is nondecreasing in a. NISP holds if $f(x|a)$ is (x, a)-affiliated (according to Milgrom and Weber (1982)). In this case, LCDFC is a local version of CISP.

1.7 Conclusion

The measurement problem—the disparity of true performance and measured performance—is at the core of many failures in incentive contracting. According to Holmstrom's standard informativeness principle, one might try to mitigate this problem by making formal contracts depend on various informative signals. This does not work, as objective and verifiable measures often contain very crude information about true performance and furthermore they incur large costs, such as legal cost, verification cost or contract-writing cost. Informal contracting (or subjective performance evaluation) could be an alternative, as it can utilize highly informative but subjective signals at little cost. However, informal contracts face the self-enforcement problem.

This paper examines when and how a new source of information mitigates contractual inefficiency given the tradeoff between informativeness and self-enforcement. I find a new informativeness principle, which is a necessary and sufficient condition of incremental information ranking. The new principle implies that statistical information creates economic value if and only if the information is sufficiently informative that the existing pass/fail criterion is refined. If the existing signal is a coarse information, an additional signal is hardly valuable. If the existing signal is a fine information, an additional signal is easily valuable. Therefore, there is an informational notion of increasing returns to scale. This notion implies that relationship is asset since a long-standing relationship is a good tool of gathering information about coworkers and this gathered existing information helps utilize new source of information. If there is a long-lasting relationship, then contractual efficiency can be easily improved by adding signals. In this case, I find an informational folk theorem, which shows that the first-best total surplus can be achieved by adding sufficiently large number of identical and independent signals. In sum, informal contracts are useful in utilizing information so that they mitigate the measurement problem.

The information structure ranking criteria are obtained given the assumption that the first-order approach (FOA) is valid. The existing FOA-justifying conditions are very strong that the set of available information structures is very small. This paper provides a remarkably weak condition (which is a local version of the existing conditions) that justifies the use of FOA in the relational contract model. Our condition is satisfied in many prominent examples (such as the additive-normal example or the additive-MLRP example) in which the existing conditions fail to be satisfied.

In this paper, I confine my focus to the case in which signals are used only informally. However, Baker et al. (1994) consider a simple hybrid contract in which a binary signal is formally used and another binary signal is informally used. See also Bernheim and Whinston (1998). As a future work, one might find an information structure ranking criterion based on a general hybrid contract $W(x, y)$ such that x is a vector of formal signals and y is a vector of informal signals.

Chapter 2

Government Bailouts, Time-Inconsistency and Managerial Punishments

2.1 Introduction

The recent 2008-09 economic crisis called for government bailouts of large-sized systemically important industrial and/or financial firms, such as General Motors or American International Group. Bailouts are considered necessary, as systemically important firms' failures could cause on the whole economy formidable negative shocks such as massive unemployment or chain-reaction bankruptcy. However, bailouts could be time-inconsistent: if managers of systemically important firms believe that their failures will be covered by governments *ex-post*, then *ex-ante* they need not exert much effort in reducing default risks—moral hazard.

Ever since Walter Bagehot notified this time-inconsistency problem in 1873, moral hazard has been a defining concern of government bailouts. The severity of moral hazard is best described by the soft budget constraint literature, which points out moral hazard as a main cause of the collapse of the communist economies.[1] The concept of soft budget constraint is not limited

[1]See Kornai (1979) who illustrates how severe is the moral hazard problem with the failure of Soviet economy. In the debate about why communist economies are inferior to market economies, Yanos Kornai attributed it to the soft-budget constraint of state-owned enterprises. Soviet states bailed out state-owned enterprises when those firms were facing severe budget deficits. Anticipating this, managers of those enterprises did not work hard.

to communist economies but is also applied broadly to capitalist economies (see Kornai et al. (2003)). For instance, the 2008-09 financial crisis witnessed devastating effects of imprudent risk management on the globally connected financial and real economies.

Accordingly, one body of literature seeks commitment devices that make nonintervention credible (see Schaffer (1989), Schmidt (1996) and Segal (1998)). Apparently, nonintervention restores ex-ante moral hazard only at the sacrifice of ex-post efficiency: firms may fail by bad luck and not only by moral hazard. Another body of literature is focused more on balancing ex-ante and ex-post efficiencies: especially, the central banking literature compares rule-based and discretionary bailouts (see Boot and Thakor (1993), Goodhart and Huang (1999), Freixas (1999) and Cordella and Yeyati (2003)). However, neither approach is a perfect solution, which obtains both ex-post efficiency (i.e. saving systemically important firms) and ex-ante efficiency (i.e. resolving moral hazard).

However, there is a simple and seemingly perfect solution—*saving firms while punishing managers*. On the one hand, the ex-post efficiency is obtained since systemically important firms are saved. On the other hand, one might guess that the ex-ante moral hazard will be restored since managers will do their best to reduce default risks in order to avoid future punishments. A body of literature advocates this idea. Aghion et al. (1999) mention, without explicit analysis, that bailouts conditional on managerial replacement can resolve moral hazard. Acharya and Yorulmazer (2008) show that if governments save a bank but take some shares off (as punishments) from *owners*, the

Consequently, productivity plummeted and thereby the communist economy failed. Kornai called this phenomenon the *soft budget constraint syndrome*.

decision makers in their model, then owners' moral hazard can be resolved. Bernardo et al. (2011) explicitly confirm that bailouts conditional on managerial replacement induce higher incentives than unconditional bailouts. In the real world, bailouts contingent on managerial punishments have been widely used. During 2008-09 economic crisis, Obama administration set a $500,000 cap on bonuses for managers of rescued financial institutions.[2] Similar measures have been taken in many other countries.[3] As the public anger against bailouts of private companies rises, such punishment measures become increasingly popular.[4] During and after the recent 2008-09 financial crisis, regulators, market participants and political leaders (especially in the G20 summit held right after the crisis) increasingly called for direct government regulation on managerial compensation scheme as attempts to reduce extraordinarily high pay level or to mitigate severe managerial moral hazard.[5]

This seemingly perfect solution works as expected if ownership and management are not separated. The existing bailout literature mainly focuses on the interaction between firms and governments but not the interaction between shareholders and managers given government intervention.[6] That is,

[2]Relatedly, the United States Troubled Asset Relief Program prohibited rescued financial firms from paying bonuses to top managers during the 2008-09 economic crisis. See the United States Emergency Economic Stabilization Act, Section 111.

[3]Brown and Dinc (2005) examine the bank bailouts data of 21 major emerging markets during the period of 1994-2000, and find that managements of 18 out of 20 bailed-out banks were ousted by governments. In South Korea, columnists claimed that the government bailout of Kia Motors, a large car maker, during the 1997-98 East Asian currency crisis would not cause a severe moral hazard problem since the owner, who had essentially managed the firm, was forced to resign. See Shin and Chang (2003).

[4]For instance, the AIG insurance corporation ignited the public outrage and strong political reactions by announcing its plan to pay about $450 million bonuses to employees in the financial service division, who have been blamed for their imprudent investments on financial derivatives, which called for the greatest ever government bailout.

[5]See section 6.2 of Brunnermeier et al. (2009).

[6]Exceptions are Chari and Kehoe (2010) and Bernardo et al. (2011). Chari and Kehoe

within-firm contracting is neglected, which could be problematic since ownership and management of systemically important firms are typically separated.

In my paper, I counter the seemingly perfect solution by explicitly analyzing within-firm contracting between shareholders and management. In particular, I show that managerial punishments could exacerbate rather than mitigate moral hazard because they distort contract design depending on corporate governance structures and punishment measures.

To get an intuition of this main result, suppose the punishment measure is Obama's pay cut, which reduces the CEO's pay on the downside. In response, shareholders may adjust the CEO's pay on the upside. The difference between this upside and downside pays is a bonus, which determines the CEO's incentives. The optimal adjustment depends on the corporate governance structure. In the strong governance case in which there are a few controlling shareholders and remaining minority shareholders, the controlling shareholders have a significant influence over the management. So they can manipulate earnings after its initial realization. If the managerial pay on the upside is too much higher than that on the downside, they will sabotage earnings in order to avoid paying the high upside pay. Therefore, only reasonable difference between upside and downside pays (i.e. bonus) is sustainable. So shareholders will reduce the upside pay in response to the pay cut on the downside. In the end, the bonus amount is unchanged and hence I have the invariance result that managerial punishments have no effect on equilibrium

(2010) explicitly consider within-firm contracting. A key difference with our paper is that the fate of firm and that of manager cannot be separated under their setting, since they assume that the only way to dismiss manager is being bankrupt, while the separation of the two fates is the core idea of bailouts conditional on managerial punishments. Bernardo et al. (2011) will be reviewed momentarily.

incentives. This illustration uses a simple binary output model. If one considers more general models, as I will do in the following, equilibrium incentives can strictly be worsen.

In contrast, if governance is weak, that is, shares are well distributed, then very high bonus can be sustained. In this case, the upside pay either increases or decreases depending on shareholders' minimum required return. Shareholders are investors so that they require some minimum amount of return. If this return is low, shareholders can pay much for the CEO, and hence, raise the upside pay. Moral hazard problem is thus mitigated. If this required return is high, shareholders are unable to raise the upside pay. More hazard is then unaffected or even exacerbated.

In general, managerial punishments (or pay regulations) are not very effective in resolving moral hazard since they (as new constraints to the contract design problem) reduce shareholders' ability to motivate management while make no effect on their willingness to motivate (as managers are punished but not shareholders). In the remainder of this paper, I consider several extensions to check the robustness of this claim. I also discuss whether shareholder or debtholder punishments could be effective alternatives.

The organization of the paper is as follows. In section 2.2, I lay out a model and show that managerial punishments, in particular pay restrictions, (weakly) exacerbate moral hazard in the strong governance case. Section 2.3 considers various extensions such as the weak governance case, managerial replacement (as an alternative form of managerial punishments), subjective performance evaluation, bargaining power and risk aversion. I find that the inefficiency of managerial punishments can be robust to these extensions. In section 2.4, as an alternative solution, I discuss the effectiveness of shareholder or debtholder punishments. Section 2.5 draws a conclusion.

2.2 The Model

There are three players, government, shareholder (she) and manager (he), and three dates $t = 0, 1, 2$. The firm, which consists of the shareholder and the manager, yields outputs y_1 and y_2 at date-1 and date-2, respectively. There is no production at date-0. I assume that there is no time-discounting without loss of generality. The timeline is as follows.

At date-0, the firm has a debt $D > 0$ (net of cash), which is due on the end of date-1. The previous macroeconomic, industrial, or financial shocks randomly determine the firm's cash and hence the debt D net of cash. The government chooses and commits to one from three bailout regimes: NB (**N**o **B**ailout), B (Simple **B**ailout), and BP (**B**ailout with a **P**ay Restriction). Under NB, the government commits not to save the firm. Under B, it could save the firm in distress. Under BP, the government could save the firm in distress but requires punishing the manager by restricting managerial compensation. An example of such pay restrictions is Obama's $500,000 cap on bonuses for rescued firms' managers.[7]

At date-1, the shareholder offers a long-term contract $\{w_1(y_1), w_2(y_1, y_2)\}$, where $w_1(\cdot)$ and $w_2(\cdot, \cdot)$ are date-1 and date-2 wages. Output y_t is pub-

[7] If pay restrictions are severe enough, distressed firms may want to decline bailout offers. In principle, firms have the option to decline bailout offers. In practice, however, governments have informal authorities to force firms to be bailed out. A related anecdote was reported during the United States financial sector bailout in 2008. On Oct. 13, 2008, the chief executives of the nine largest banks in the United States were forced to sign a government bailout plan by the Treasury Secretary Henry Paulson. One of the executives said it was a 'take-it-or-take-it' offer. See "Drama behind a $250 billion banking deal", *New York Times*, Oct. 13, 2008

licly observable and verifiable. Given a contract, the manager chooses the level of a privately-observable effort $a_1 \in [0, \infty)$, which reduces the firm's default probability and increases the firm's expected output. The cost of effort $c(a_1)$ is increasing and convex. This effort and some random shocks (such as macroeconomic, industrial or financial shocks) determine the date-1 output $y_1 \in Y \subset [0, \infty)$ according to a probability distribution function $F(y_1|a_1)$. After y_1 realizes, the firm needs to repay the debt D to continue operation at date-2. I shall explain the detail of repayment procedure momentarily.

At date-2, the manager chooses a privately-observable effort $a_2 \in \{h, l\}$, where $h > l$. After the effort choice, the date-2 output $y_2 \in \{g, m\}$ realizes, where $g > m \geq D$. How a_2 affects y_2 will be explained shortly.

Efforts

date-1 effort choice is the focus of this paper. Let $f(y_1|a_1) > 0$ be the density function of the distribution $F(y_1|a_1)$. Let $f_a(y_1|a_1)$ be the partial derivative of $f(y_1|a_1)$ with respect to a_1. In order to model how date-1 effort a_1 affects date-1 output y_1, I assume that the probability density function $f(y_1|a_1)$ satisfies the following monotone likelihood ratio property (MLRP):

Assumption 2.1. $\frac{f_a(y_1|a_1)}{f(y_1|a_1)}$ *is strictly increasing in* y_1 $\forall a_1$

MLRP implies two things. First, the default probability $\mathbb{P}(y_1 \leq D|a_1)$ is decreasing in a_1. Second, the expected output $\mathbb{E}[y_1|a_1]$ is increasing in a_1. That is, low a_1 means moral hazard, as low a_1 implies high-risk but low-return.

I do not interpret a_1 as the hours spent in the office, as most CEOs work for very long hours. Instead, I interpret a_1 as prudent risk management activities (such as monitoring investment projects, restructuring portfolio or avoiding excessive risk-taking) that reduce the firm's default risk without sacrificing return too much. Similar interpretations are used in Aghion et al. (1999), Tirole (2010), Dell'Ariccia and Ratnovski (2013), Benabou and Tirole (2013), etc.

As an extension, one might also consider that a_1 can contribute to debt restructuring. The outstanding debt D is a short-term debt, which should be paid at date-1. If the manager undertakes negotiation with the debtholders over the term structure, the firm can increase the maturity. Debt negotiation is typically very painful since there are typically very different types (like banks, institutional lenders, individuals, government, etc.) of many debtholders. The diversity of interests, seniority, covenants and transaction cost make debt restructuring very difficult. Let $\delta(a_1)$ denote the portion of restructured debt. I assume that $\delta(a_1)$ is increasing in a_1. Then, the default probability $\mathbb{P}(y_1 < D - \delta(a_1)|a_1)$ is still decreasing in a_1. Since the effect of a_1 on the default probability is unchanged (and hence the main result is unaffected), I simply assume that debt restructuring is impossible for the following analysis.

Since date-2 effort a_2 is not the focus of this paper, I simplify the analysis by assuming that a_2 is binary: it is either "behaving" (or high effort h) or "misbehaving" (or low effort l). I regard misbehaving any kind of improper but privately attractive management activities: showing favoritism in human resources management, involving less with profit-making but more with unrelated activities (such as business party or political involvement), etc.[8] If misbehaves, he can enjoy some private benefit $B > 0$. The date-2 output y_2

[8]These examples are borrowed from Tirole (2010), which provides much fuller interpretations of behaving/misbehaving.

is then either "good" g or "mediocre" m, where $g > m \geq D$.[9] If the manager behaves, the probability of good output is $p_h \equiv \mathbb{P}(y_2 = g|a_2 = h)$; if he misbehaves, it is p_l, where $\Delta p \equiv p_h - p_l > 0$.

I assume that behaving is much more profitable than misbehaving:

Assumption 2.2. $\mathbb{E}[y_2|h] \gg \mathbb{E}[y_2|l] + B \quad \Leftrightarrow \quad g - m \gg \frac{B}{\Delta p}$

where $\mathbb{E}$ is expectation. Based on this assumption, I am interested in equilibrium where $a_2 = h$ is chosen.

Repayment and bailouts

At the end of date-1, the firm has to repay the debt D by using the cash on hand y_1. The firm may also borrow from the capital market. If the financial system is perfect, the firm can borrow up to the future prospect $\mathbb{E}[y_2|h]$. But due to various sources of financial market inefficiency (such as debt-overhang problem, transaction cost or renegotiation problem), the firm borrows only a fraction of its future income. This is particularly the case when the financial market is in systemic risk. In the remainder of this paper, I assume that the firm cannot borrow anything without loss of generality. Thus, in the no-bailout regime NB, the firm's long-term value $v^N(y_1)$ is given by

$$v^N(y_1) = \mathit{1}_{\{y_1 \geq D\}}(y_1 + \mathbb{E}[y_2|h] - D) \tag{2.1}$$

where $\mathit{1}_{\{\}}$ is the indicator function.

[9] Since $m \geq D$, there is no date-2 default.

Since the firm is systemically important, its bankruptcy could cause a massive negative shock to the overall economy, and thereby, the social welfare is reduced by $E > 0$. Under the bailout regimes B and BP thus the government may inject a relief fund of amount k at the cost of bailout $\psi(k)$ such that $\psi(0) = 0$, $\psi'(k) \geq 0 \ \forall k$. This cost may reflect the opportunity cost of the fund or the political burden. Relatedly, Brown and Dinc (2005) find that government intervention is much less likely before election than after election, by examining distressed large private banks in 21 emerging countries in 1990s. Bailouts of many large firms with budget could also increase sovereign credit risk. Acharya et al. (2014) show empirically that 2008-09 European government bailouts increased sovereign credit default risk.

If the government chooses to save the firm, then since bailouts are costly, the government will lend only the tight money $D - y_1$ at the market interest rate.[10] Then, the firm continues to date-2, earns y_2, repays the relief fund and then takes the remainder $y_2 - (D - y_1)$. Consequently, bailouts leave the firm a rent of $\mathbb{E}[y_2|h] - (D - y_1) > 0$ in expectation, which could not have been obtained without government intervention. Thus, the firm's long-term value $v(y_1)$ at date-1 is

$$v(y_1) = y_1 + \mathbb{E}[y_2|h] - D \quad \forall y_1 \tag{2.2}$$

[10]The government may lend more than the tight money at a preferential interest rate or based on unsecured collateral, so that the rescued firms get some direct subsidies from bailouts. Paul Krugman claimed that the Geithner plan—a U.S. bailout policy during the economic crisis 2008-09—essentially ended up with subsidizing the investors of the rescued financial firms. He attributed the reason to that the government provided relief funds mostly as of non-recourse loans, and hence, the investors is effectively given a put-option. See "Geithner plan arithmetic", *New York Times*, March 23, 2009. Relatedly, the Federal Reserve issued a non-recourse debt of $30 billion to JP Morgan Chase to help it purchase Bear Sterns on March 16, 2008. The collateral of the non-recourse loan is Bear Sterns' less liquid assets. Therefore, the Federal Reserve should absorb the loss, if the collateralized assets' values fall.

I assume that bailouts are undesirable due to the high bailout cost $\psi(D-y_1)$ but necessary in order to prevent the greater worse, the negative spillover effect E. That is,

$$-E + y_1 < y_1 + \mathbb{E}[y_2|h] - D + D - \psi(D - y_1) < 0 \quad \forall y_1 < D \tag{2.3}$$

where the first two terms sum to the social value without bailouts, the next five terms sum to the social value with bailouts (i.e. the sum of firm value, debt value and negative of bailout cost). If the firm is solvent ($y_1 \geq D$), the government is assumed to be unable to intervene. Then, the social value is simply equal to the firm value $v(y_1)$. Therefore, a (constant-adjusted) social value $s(y_1)$ is given by[11]

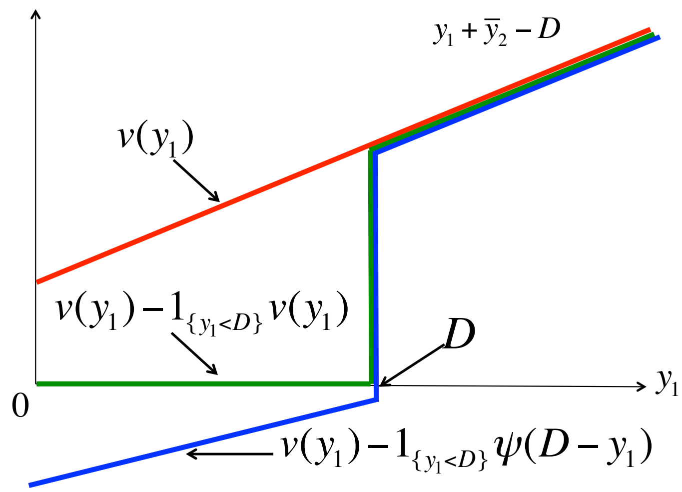
$$s(y_1) = v(y_1) - \mathit{1}_{\{y_1 < D\}} \cdot \psi(D - y_1) \tag{2.4}$$

See Figure 2.1, which illustrates firm value with bailouts $v(y_1)$, firm value without bailouts $v^N(y_1) = v(y_1) - \mathit{1}_{\{y_1<D\}} v(y_1)$ and (constant-adjusted) social value with bailouts $s(y_1)$.

The main objective of this model is to find how bailout regimes B and BP affect equilibrium date-1 efforts. Equilibrium date-1 efforts are induced by optimal incentive contracts, as the shareholder or the government cannot observe effort choices. Later, I solve for optimal incentive contracts and corresponding equilibrium efforts. In the following, I consider benchmark cases in which the shareholder or the government observes date-1 effort choices and force the manager to choose the levels what they want.

[11] I can ignore the constant D in determining socially efficient effort level a_1^{**}, as the constant term has no incentive effect.

Figure 2.1: Firm value with and without bailouts and social value with bailouts (where $\bar{y}_2 = \mathbb{E}[y_2|h]$)



Benchmark levels of date-1 efforts

If the shareholder observes a_1, she will choose total surplus maximizing efforts. Let $a_1^{N*} > 0$ be the total surplus maximizing effort under NB, which is characterized via

$$\int v^N(y_1) f_a(y_1|a_1) dy_1 = Cov\left(v^N(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = c'(a_1) \;\; at \;\; a_1 = a_1^{N*} \tag{2.5}$$

[12]

where the covariance term $Cov(\cdot)$ represents the marginal incentive. Note that $a_1^{N*} > 0$ since the marginal incentive is positive at $a_1 = 0$. Similarly, the total surplus maximizing effort $a_1^* > 0$ under B or BP is characterized via

$$Cov\left(v(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = c'(a_1) \;\; at \;\; a_1 = a_1^* \tag{2.6}$$

[12] I have the covariance representation since the expectation of $\frac{f_a(y_1|a_1)}{f(y_1|a_1)}$ is zero.

If the government observes a_1, it will choose the social welfare maximizing effort a_1^{**}, which is characterized via

$$Cov\left(s(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = c'(a_1) \ \ at \ a_1 = a_1^{**} \tag{2.7}$$

I compare the three ideal efforts. See Figure 2.1. Compare the social value $s(y_1)$ to the firm value without bailouts. The social value is equivalent to the firm value on the upside but is less than the firm value on the downside since bailouts are undesirably costly. Thus, the government's marginal incentive to increase the output y_1 is higher than the shareholder's. That is, I have $a_1^{**} > a_1^{N*}$. Compare now the firm value without bailouts to the firm value with bailouts $v(y_1)$. Since bailouts cover the downside risk, the shareholder's marginal incentive to increase the output is reduced if she expects bailouts. That is, I have $a_1^{N*} > a_1^{*}$.

To formalize this reasoning, consider the following related technical assumption and lemma.

Assumption 2.3. *$F(y_1|a_1)$ is strictly concave in a_1 $\forall y_1$.*

The convexity of distribution function condition (CDFC) above is a standard technical assumption by which incentive problems become tractable.

Lemma 2.1. *Let $g(y_1)$ be a nontrivially increasing function, that is, $g(y_1)$ is nondecreasing on Y and strictly increasing on a set of positive measure. Suppose Assumption 2.1 and 2.3. Then, $\int_0^\infty g(y_1)f(y_1|a_1)dy_1 - c(a_1)$ is inverse U-shaped and is strictly concave in a_1.*

Proof: Through integration by parts, I have

$$\mathbb{E}[g(y_1)|a] = g(0) + \int_0^\infty [1 - F(y_1|a_1)]dg(y_1)$$

(Strict) MLRP in Assumption 2.1 implies (strict) first-order stochastic dominance, that is, $F_a(y_1|a_1) < 0\ \forall a$ and $\forall y_1 \in Y\backslash\{0\}$. Then, since $dg(y_1) \geq 0$ on Y and $dg(y_1) > 0$ on a set of positive measure, I have $\frac{d\mathbb{E}[g(y_1)|a]}{da} = -\int_0^\infty F_a(y_1|a_1)dg(y_1) > 0$. (Strict) CDFC in Assumption 2.3 means $F_{aa}(y_1|a_1) > 0$. Thus, I have $\frac{d^2\mathbb{E}[g(y_1)|a]}{da^2} = -\int_0^\infty F_{aa}(y_1|a_1)dg(y) < 0$. ■

Lemma 2.1 implies that two optimal efforts can be compared by corresponding marginal incentives (the covariance terms): for two nontrivially increasing functions $m(y_1)$ and $n(y_1)$, let a_1^m and a_1^n be maxima of $\int_0^\infty m(y_1)f(y_1|a_1)dy_1 - c(a_1)$ and $\int_0^\infty n(y_1)f(y_1|a_1)dy_1 - c(a_1)$ with respect to a_1, respectively. Then, $a_1^m > a_1^n$ if $Cov\left(m(y_1) - n(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) > 0$ at either a_1^m or a_1^n.

By Assumption 2.1, there exists $\hat{y}(a_1) \in Y$ such that $f_a(y_1|a_1) > 0$ if $y > \hat{y}(a_1)$ and $f_a(y_1|a_1) < 0$ if $y < \hat{y}(a_1)$. Thus, $\hat{y}(a_1)$ can be regarded a cutoff of good and bad performance. Note that D is the critical point of bailouts. If bailouts take place when the performance is sufficiently bad, then one can regard D a cutoff of critical failure while $\hat{y}(a_1)$ a cutoff of non-critical failure. Hence, I can reasonably assume that:

Assumption 2.4. *$D \leq \hat{y}(a_1^{N*})$, where $\hat{y}(a_1)$ is such that $f_a(\hat{y}(a_1)|a_1) = 0$.*

Then, at $a_1 = a_1^{N*}$, I have

$$Cov\left(s(y_1) - v^N(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = Cov\left(s(y_1) + D - v^N(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right)$$

$$= \int_0^D \underbrace{\{y_1 + \mathbb{E}[y_2|h] - \psi(D - y_1)\}}_{<0 \text{ by } (2.3)} \underbrace{f_a(y_1|a_1)}_{<0} dy_1 > 0$$

and

$$Cov\left(v^N(y_1) - v(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = -\int_0^D v(y_1) \underbrace{f_a(y_1|a_1)}_{<0} dy_1 > 0$$

Then, by Lemma 2.1, it follows that

$$a_1^* < a_1^{N*} < a_1^{**} \tag{2.8}$$

Contracting under bailouts without a pay restriction

Next I consider a more realistic case in which efforts can be induced only by incentive contracts. I first construct the contracting program under B, the bailout regime without any pay restriction. At the beginning of date-1, the shareholder offers a long-term contract $\{w_1(y_1), w_2(y_1, y_2)\}$ in order to motivate the manager. Obviously, not all contracts are available. In the following, I consider constraints that available contracts must satisfy.

Incentive constraints: Given a long-term contract, the manager chooses efforts at date-1 and date-2. I first consider the date-2 incentive constraint. For a given y_1, the manager chooses to behave if the following date-2 incentive constraint is satisfied:

$$\mathbb{E}[w_2(y_1, y_2)|h] \geq \mathbb{E}[w_2(y_1, y_2)|l] + B \quad \Leftrightarrow \quad w_2(y_1, g) - w_2(y_1, m) \geq \frac{B}{\Delta p} \tag{IC2}$$

Second, I consider the date-1 incentive constraint. A long term contract induces a long-term compensation $w(y_1) \equiv w_1(y_1) + \mathbb{E}[w_2(y_1, y_2)|h]$. At date-1, the manager chooses an optimal effort a_1 that maximizes his expected payoff:

$$a_1 \in \max_{\hat{a}} \int w(y_1) f(y_1|\hat{a}) dy_1 - c(\hat{a}) \tag{IC1}$$

The manager's long-term payoff is then given by

$$u \equiv \int w(y_1) f(y_1|a_1) dy_1 - c(a_1) \tag{U}$$

Participation constraints: If the manager leaves the firm, he will get a new job or begin a start-up. Let $\bar{u}$ be the manager's reservation payoff from these outside options. As an investor, the shareholder requires a minimum return from the investment on the firm. Let $\bar{\pi}$ be the shareholder's reservation payoff from other investment opportunities. Acceptable long-term contracts must satisfy the following participation constraints:

$$u \geq \bar{u} \tag{PC-M}$$

$$\pi \equiv \int [v(y_1) - w(y_1)] f(y_1|a_1) dy_1 \geq \bar{\pi} \tag{PC-S}$$

Boundary constraints: There are bounds on contracts, as the following boundary constraints describe:

$$A \equiv p_h \frac{B}{\Delta p} \leq w(y_1) \leq v(y_1) \quad \forall y_1{}^{13} \tag{BC}$$

The upper bound arises due to the shareholder's limited liability constraint:[14] the shareholder cannot credibly commit to pay the manager more than the firm

[13] One might ask why not the maximum feasible compensation is $y_1 + \mathbb{E}[y_2|h]$ rather than $y_1 + \mathbb{E}[y_2|h] - D$. I assume that the debtholders can set covenants that secure debt repayment. Accordingly, the shareholder will be prevented to pay the manager more than the firm value.

[14] In fact, limited liability constraints must be defined for each date $t = 1, 2$. If the firm is solvent ($y_1 \geq D$), its available income is $y_1 - D$ at date-1 and y_2 at date-2. Then, (per-period) limited liability constraints are given by:

$$0 \leq w_1(y_1) \leq y_1 - D \quad \forall y_1 \geq D$$

$$0 \leq w_2(y_1, y_2) \leq y_2 \quad \forall y_2 \in \{g, m\}, \text{ for each given } y_1 \geq D \tag{2.9}$$

(Note: In the limited liability constraints above, the upper bound of date-2 compensation is the date-2 income. Instead, one might consider that the upper bound is the date-2 wealth, that is, the sum of date-1 income after paying the date-1 compensation and the date-2 income. Main results are robust to this alternative consideration.)

If the firm is insolvent ($y_1 < D$), it borrows the tight money $D - y_1$ from the government. Thus, it has no remaining income at date-1. At date-2, the firm's available income is $y_2 - (D - y_1)$. The limited liability constraints are then given by

$$0 \leq w_1(y_1) \leq 0 \quad \forall y_1 < D$$

value. The lower bound arises due to the manager's limited liability constraint and the date-2 incentive constraint (IC2):

$$w(y_1) = w_1(y_1) + p_h[w_2(y_1, g) - w_2(y_1, m)] + w_2(y_1, m) \geq 0 + p_h \frac{B}{\Delta p} + 0$$

To motivate the manager at date-2, there must be a sufficient wedge (i.e. at least $\frac{B}{\Delta p}$) in the date-2 wages. In the beginning of date-1 thus the manager can guarantee the expected wedge $A \equiv p_h \frac{B}{\Delta p}$. I call A the agency rent, as it is given to the manager because his date-2 behavior is privately observable. The long term wage $w(y_1)$ should then be no less than the agency rent.

In the following, I consider an optional constraint, whose presence plays an important role throughout this paper.

Monotonicity constraint: strong/weak governance

The shareholder's ex-post payoff is $v(y_1) - w(y_1)$. The monotonicity constraint states that this ex-post payoff is nondecreasing in y_1:

$$v(y_1) - w(y_1) \textit{ is nondecreasing in } y_1 \ \forall y_1 \tag{MC}$$

Suppose a contract $w(y_1)$ does not satisfy (MC). For $y' < y''$, let $w(y') > w(y'')$. Then, the shareholder is worse off as the firm's output increases from

---

$$0 \leq w_2(y_1, y_2) \leq y_2 - (D - y_1) \quad \forall y_2 \in \{g, m\}, \text{ for each given } y_1 < D \tag{2.10}$$

If a long-term contract $\{w_1(y_1), w_2(y_1, y_2)\}$ satisfies (IC2), (2.9), and (2.10), then it induces a long-term compensation $w(y_1) \equiv w_1(y_1) + \mathbb{E}[w_2(y_1, y_2)|h]$, which satisfies (BC). Conversely, by Assumption 2.2, for any long-term compensation $w(y_1)$ that satisfies (BC), there exists a long-term contract $\{w_1(y_1), w_2(y_1, y_2)\}$ that satisfies (IC2), (2.9), (2.10), and induces $w(y_1)$.

y' to y'', which is abnormal in practice. In this case, if the realization of y_1 is y'', the shareholder is strictly better off by sabotaging output by $y'' - y'$. But then, $w(y_1)$ essentially satisfies (MC).

The shareholder may be able to sabotage output if she is deeply involved with management, which will be the case in practice when there are a few controlling shareholders and remaining minority shareholders. I call this case *strong governance*.[15] In contrast, the shareholder is not able to sabotage output if management is entirely delegated to the manager, which will be the case in practice when shares are well distributed. I call this case *weak governance*. The main body of this paper focuses on the strong governance case. In an extension, I consider the weak governance case.

Contracting program: All of the constraints above imply that a long-term compensation $w(y_1)$ is what really matters for date-1 contracting. In this regard, I simply call $w(y_1)$ a long-term contract. I assume that the labor market for top managers is competitive. The shareholder thus maximizes management's payoff in order to hire a top talented manager provided that her minimum required return is satisfied. That is, the manager has full bargaining power. In an extension, however, I show that the main result is robust to the allocation of bargaining power. Thus, the date-1 long-term contracting

[15]Note that (MC) is the date-1 monotonicity constraint. Under the strong governance, I may also need to consider the date-2 monotonicty constraint, which requires

$$g - w_2(y_1, g) \geq m - w_2(y_1, m) \quad \text{for each } y_1 \tag{MC2}$$

But (MC2) is redundant: for any $w(y_1)$ that satisfies (BC), there exists $\{w_2(y_1, y_2)\}$ that satisfies (MC2), (IC2), (2.9), and (2.10).

program is given by:

$$\max_{a\in[0,\infty),\ w:Y\to\mathbb{R}} u \quad \text{subject to (U), (IC1), (PC-M), (PC-S), (BC), and (MC)} \tag{2.11}$$

The contracting program (2.11) is typically intractable, as it is subject to an arbitrary maximization constraint (IC1). Since the constraint is not an inequality condition, usual techniques (such as Lagrange or Kuhn-Tucker method) cannot be used. The contracting program becomes tractable if I replace (IC1) with the following first order condition:

$$\int w(y_1) f_a(y_1|a_1) dy_1 = Cov\left(w(y_1), \frac{f_a(y_1|a_1)}{f(y_1|a_1)}\right) = c'(a_1) \tag{FOC}$$

Given this first order approach, the set of feasible contracts can be defined via

Definition 2.1. A contract $w(y_1)$ is $(\bar{\pi}, \bar{u})$-admissible (or simply, admissible) if it induces $a_1 > 0$[16] and satisfies (FOC), (PC-M), (PC-S), (BC), and (MC).

Let $\mathbb{A}(\bar{\pi}, \bar{u})$ denote the set of $(\bar{\pi}, \bar{u})$-admissible contracts. Then, assuming that the first order approach is valid, I solve the following *relaxed* contracting program[17]

$$\max_{a\in[0,\infty),\ w(y_1)\in\mathbb{A}(\bar{\pi},\bar{u})} u \quad \text{subject to (U)} \tag{2.12}$$

The validity of the first order approach shall be checked later.

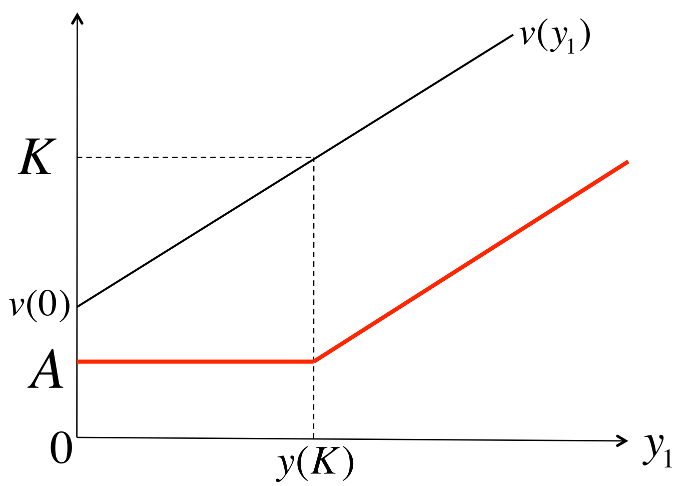
[16]That a contract induces a_1 means (FOC) is satisfied at a_1.

[17]Note that (IC1) implies (FOC) but not vice versa. Thus, the choice set with (FOC) is greater than the true choice set. However, if an optimal contract $w^*(y_1)$, which solves the contracting program subject to (FOC) is nontrivially increasing (and induces a positive effort), then (FOC) implies (IC1) by Lemma 2.1. Then, $w^*(y_1)$ is a true optimal contract. I later check that $w^*(y_1)$ is nontrivially increasing.

In the strong governance case, the following type of contracts is particularly important:

Definition 2.2. (i) A contract $w(y_1)$ is called a stock option if it equals $[v(y_1) - K]^+ \equiv \max\{v(y_1) - K, 0\}$ with a strike price $K \geq 0$. (ii) A contract $w(y_1)$ is called an augmented stock option if it is the sum of a stock option $[v(y_1) - K]^+$ and a fixed salary, which is equal to $p_h \frac{B}{\Delta p}$.

Figure 2.2: An augmented stock option where $y(K) \equiv \mathit{1}_{\{v(0)<K\}} v^{-1}(K)$



Innes (1990) considers a one-period lending relationship between entrepreneur and investor with limited liability and monotonicity constraints under the absence of date-2 effort choice, the liquidity shortage problem, and government intervention. He shows that optimal lending contracts have the form of stock option.[18] The following lemma and corollary extend his result to

[18]In fact, Innes (1990) shows that debt contracts are optimal compensation schemes for the

this two-period contracting model, which incorporates necessary instruments regarding government bailouts.

Lemma 2.2. *Suppose there is an admissible contract that is not an augmented stock option. Let* $a_1 > 0$ *be the effort the contract induces. Then, there is an admissible augmented stock option, which induces* $a_1(K) > a_1$ *and makes both the shareholder and manager strictly better off.*

Proof: See Lemma 1 and 2, Property (2), and Equation (3.5) of Innes (1990).

Corollary 2.1. *Suppose there is a contract* $w^*(y_1)$ *that solves the relaxed contracting program (2.12) and induces an effort* $a_1 > 0$*. Then, (i)* $w^*(y_1)$ *is an augmented stock option and (ii)* $(w^*(y_1), a_1)$ *is a solution to the exact contracting program (2.11).*

Proof: (i) If $w^*(y_1)$ is not an augmented stock option, Lemma 2.2 implies there is an admissible augmented stock option that outperforms $w^*(y_1)$, which is a contradiction. (ii) Augmented stock options are nontrivially increasing. Lemma 2.1 then implies that (IC1) is satisfied. $w^*(y_1)$ is thus a solution to (2.11). ■

The augmented stock options are optimal because they are most powerful feasible contracts. To see this briefly, note that these contracts pay nothing

---

investor. Thus optimal compensation schemes for the entrepreneur are equity contracts. For the entrepreneur, equity contracts are having the form in Definition 2.2 since the ownership is transferred to the investor when the entrepreneur cannot meet the target K. In the employment relationship, as in this paper, the same form can be interpreted as augmented stock options.

on the downside (i.e. $y_1 < y(K) \equiv \mathit{1}_{\{v(0)<K\}} v^{-1}(K)$) but pay the most on the upside. The manager thus faces the greatest possible incentives.

In principle, the contracting program (2.11) is complicated as it solves for a *function* $w(y_1)$. Fortunately, it can be greatly simplified by the optimality of an augmented stock option, which is completely characterized by a *scalar* K. Therefore, everything relevant is determined by the level of K. In the following, I examine how the strike price K affects incentives and welfares of the shareholder and manager.

First, the date-1 incentives decrease in the strike price K. To see this briefly, note that a stock option provides insurance on the downside ($y_1 < y(K)$) and incentive on the upside ($y_1 \geq y(K)$). If the strike price K is lower than $v(0)$, there is no insurance but only incentive, and hence, the manager chooses the total surplus maximizing effort a_1^*. As the strike price rises, the downside insurance increases while the upside incentive decreases, and thereby, he chooses lower effort. To see this formally, note first that an augmented stock option induces a positive effort $a_1(K)$, which is uniquely characterized by the following incentive constraint (FOC):[19]

$$\int_{y(K)}^{\infty} [v(y_1) - K] f_a(y_1|a_1) dy_1 = c'(a_1) \quad at \;\; a_1 = a_1(K) > 0 \qquad (2.13)$$

If $K \leq v(0)$, then $y(K) = 0$. The augmented stock option then reduces to $v(y_1)$ minus some constant, which means the manager becomes a residual claimant. He will then choose the total surplus maximizing effort a_1^* (see (2.6)

[19]Note that an augmented stock option is nontrivially increasing. Lemma 2.1 implies then $a_1(K)$ exists uniquely. Suppose $a_1(K) = 0$ for some $K \geq 0$. The marginal cost at $a_1 = 0$ is 0. The marginal revenue at $a_1 = 0$ is $\int_{y(K)}^{\infty} [v(y_1) - K] f_a(y_1|0) dy_1 = -\int_{y(K)}^{\infty} F_a(y_1|0) dv(y_1) > 0$ through integration by parts, a contradiction to (FOC).

and (2.13)). If $K > v(0)$, the higher the strike price K, the lower the induced effort $a_1(K)$: (2.13) and the implicit function theorem implies

$$\text{sign}\{a_1'(K)\} = \text{sign}\{F_a(y(K)|a_1(K))\} \underbrace{< 0}_{\text{by Assumption 2.1}} \quad \forall K > v(0) \tag{2.14}$$

Second, the manager is better off while the shareholder is worse off as the strike price K rises. This is simply because the upside incentive pay decreases in K. Thus, the manager receives less (and the shareholder pays less). To see this formally, let $\pi(K)$ and $u(K)$ be such that

$$\pi(K) \equiv \int_0^{y(K)} v(y_1) f(y_1|a_1(K))dy_1 + \int_{y(K)}^{\infty} K f(y_1|a_1(K))dy_1 \tag{2.15}$$

$$u(K) \equiv \int_{y(K)}^{\infty} [v(y_1) - K] f(y_1|a_1(K))dy_1 - c(a_1(K)) \tag{2.16}$$

Then, the shareholder and manager' ex-ante payoff given $a_1(K)$ are $\pi(K) - A$ and $u(K) + A$, respectively. The envelope theorem implies:

$$u'(K) = -[1 - F(y(K)|a_1(K))] < 0 \quad \forall K \geq 0 \tag{2.17}$$

Recall that the contracting program seeks to maximize the manager's payoff subject to the shareholder's participation. The shareholder thus needs to lower the strike price as much as possible in order to maximize the manager's payoff. By (2.14) then the contracting program (2.11) is simplified to

$$\max_{K \geq 0} \; a_1(K) \quad \text{subject to} \quad \pi(K) \geq \bar{\pi} + A \text{ and } u(K) \geq \bar{u} - A \tag{2.18}$$

Note that (2.18) is a key observation in this section: the manager's payoff maximization program is equivalent to the incentive maximization program. Since $a_1(K) \leq a_1^* < a_1^{**} \; \forall K$, the government also wants to maximize effort. Therefore, the shareholder and government's incentives are aligned.

Hitherto, I consider the contracting program under the bailout regime with no punishments (B). Now, I consider the bailout regime with an arbitrary pay restriction (which is effective if the firm is insolvent ($y_1 < D$)) as a managerial punishment. An example of such a pay restriction is the following bonus cap:

$$|w_2(y_1, g) - w_2(y_1, m)| \leq \Sigma \ \ \forall y_1 < D$$

If Σ is smaller than the agency rent A, the manager never behaves at date-2. In this section, I confine my focus to an arbitrary pay restriction, which is not inconsistent with the date-2 incentive constraint (IC2), that is, $\Sigma > A$ for the example above. (In an extension, I also consider pay restrictions that are inconsistent with the date-2 incentive constraint). Let (P) denote such a pay restriction. Then, the contracting program under the bailout regime BP is given by

$$\max_{a \in [0,\infty),\ w: Y \to \mathbb{R}} u \quad \text{subject to (U), (IC1), (PC-M), (PC-S), (BC), (MC), and (P)} \tag{2.19}$$

Note that this contracting program under BP is equivalent to the contracting program (2.11) under B except the additional pay constraint (P). Note also that the original contracting program (2.11) is equivalent to the incentive maximization program (2.18). One might then guess that the additional constraint (P) ends up reducing equilibrium incentives. This is intuitively correct but technically incorrect, as optimal contracts are not necessarily augmented stock options under the bailout regime BP due to the pay restriction (P). The following theorem is the first main result of this paper with a correct proof.

Theorem 2.1. *Let a_1^B be the equilibrium effort under the bailout regime B without a pay restriction. Let a_1^{BP} be the equilibrium effort under the bailout* *regime BP with an arbitrary pay restriction (P) (which is effective if the firm is insolvent ($y_1 < D$)), where (P) is not inconsistent with (IC2). Then,*

$$a_1^{BP} \leq a_1^B \leq a_1^* < a_1^{**}$$

Proof: Consider the following lemma first:

Lemma 2.3. *Consider the bailout regime B. Suppose there is an admissible contract. Then, there exists a unique optimal contract, which induces a unique equilibrium effort.*

Proof: By hypothesis and Lemma 2.2, there are admissible augmented stock options (with some strike prices). Then, (2.18) implies the choice set $\mathcal{K} \equiv \{K \in [0, \infty) : \pi(K) \geq \bar{\pi} + p_h \frac{B}{\Delta p}, u(K) \geq \bar{u} - p_h \frac{B}{\Delta p}\}$ is not empty. Let $\hat{K}$ denote the infimum of $\mathcal{K}$. Then, $\hat{K} \in \mathcal{K}$ since $\pi(K)$ and $u(K)$ are both continuous. That is, $\hat{K}$ is a solution to (2.18). By (2.14), $\hat{K}$ is the unique solution, which induces a unique equilibrium effort $a_1(\hat{K})$. ■

By Lemma 2.3, the bailout regime B induces a unique effort $a_1^B = a_1(K^B)$ with the optimal strike price K^B. By (2.8) and (2.13), it follows that $a_1^B = a_1(K^B) \leq a_1^* < a_1^{**}$. The optimal contract under BP, which induces a_1^{BP} is either an augmented stock option or an another contract. If the contract is an augmented stock option with a strike price K', it is feasible under B. (2.14) and (2.18) imply that $a_1^B = a(K^B) \geq a(K') = a_1^{BP}$. If the optimal contract under BP is not an augmented stock option, then there exists an admissible augmented stock option with a strike price K'', which induces $a_1(K'') > a_1^{BP}$ by Lemma 2.2. Then, by the similar logic, I have $a_1^B > a_1^{BP}$. ■

Figure 2.3: Pay restrictions exacerbate moral hazard

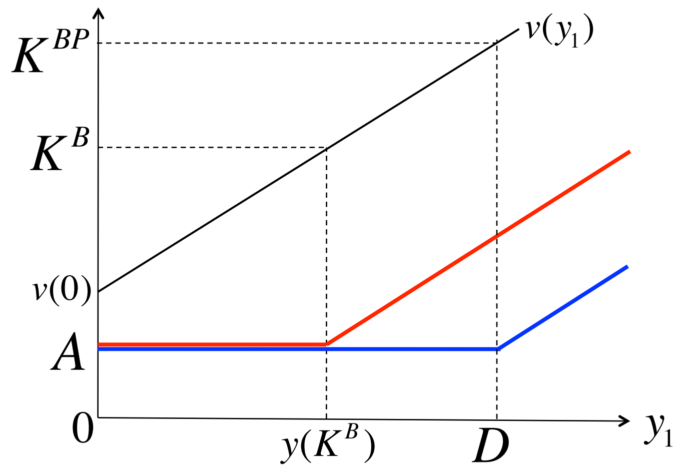


Figure 2.3 illustrates Theorem 2.1. Let K^B be the optimal strike price under the bailout regime B. Suppose that the following pay restriction is added

$$w(y_1) = A \quad \forall y_1 < D \tag{2.20}$$

That is, if the firm is insolvent, it pays the minimum level that is consistent with (IC2). Suppose also that the pay restriction is binding, that is, $D > y(K^B)$. Previously without the pay restriction, some incentives are provided in the downside ($\{y_1 : y_1 < D\}$). The pay restriction however eliminates that incentives. In response, the shareholder may wish to increase incentives on the upside. To raise incentives on the upside, the shareholder needs to increase the slope of contract on the upside. This is impossible because the slope of the contract is already at its maximum $v'(y_1)$. Greater slope violates the monotonicity constraint (MC).

Theorem 2.1 is very general: it is true for any information structure $F(y_1|a_1)$, any increasing and convex cost function $c(a_1)$, any nontrivially increasing income schedule $v(y_1)$, and any parameter values $(\bar{\pi}, \bar{u})$ provided Assumptions 2.1–2.4 hold.

Theorem 2.1 depends on two assumptions. First, the shareholder maintains strong influence over the management so that she can sabotage output and hence (MC) must be satisfied. Second, I confine my focus to pay restrictions that are not inconsistent with (IC2). In the following extensions, I relax these assumptions.

2.3 Extensions

2.3.1 Weak Governance

Under weak governance, managerial punishments mitigate moral hazard if the shareholder's required return $\bar{\pi}$ is low, whereas they exacerbate moral hazard otherwise. To illustrate this idea heuristically, see Figure 2.3. The particular punishment (2.20) eliminates incentives on the downside. To compensate for this loss in incentives (and hence to raise the manager's payoff), the shareholder wants to raise incentives on the upside. To do that, the slope of wage should be raised, which means the manager is paid more. This is possible if the shareholder affords to pay more (i.e. the required return is low), but not otherwise.

This illustration is heuristic since augmented stock options are no more optimal in the weak governance case. However, the main message is still true: the larger the two parties' contractual friction (i.e. $\bar{\pi}$), the more difficult to find a powerful contract, which provides strong incentives to the manager. The following analysis formalizes this reasoning.

The (unrelaxed) contracting program under the bailout regime B is then given by

$$\max_{a \geq 0,\ w:Y \to \mathbb{R}} u \quad \text{subject to (U), (PC-M), (PC-S), (BC), and (IC1)} \tag{2.21}$$

Previously, I consider a *relaxed* contracting program in which (IC1) is replaced by (FOC) and then checked its validity. In this section, I consider a *doubly-relaxed* contracting program in which (IC1) is replaced by the following inequality constraint:

$$\int w(y_1) f_a(y_1|a_1) dy_1 \geq c'(a_1) \tag{RFOC}$$

This is the relaxed first order approach (RFOA), proposed by Rogerson (1985), which is quite a convenient technique since the sign of the Lagrange multiplier of (RFOC) cannot be negative. Note that (IC1) implies (RFOC) (if the equilibrium effort is positive) but not vice versa, and hence, the doubly-relaxed program has a greater choice set than the exact program. Thus, if an optimal contract to the doubly-relaxed program is nontrivially increasing, then by Lemma 2.1 the contract satisfies (IC1) and thus is a solution to the exact program. Thus, a contract is admissible if it satisfies (RFOC), (PC-M), (PC-S), (BC), and induces a positive effort.

Innes (1990) characterizes optimal contracts under the absence of government intervention. His result can be directly extended to the doubly-relaxed contracting program. The following lemma provides a partial characterization of optimal contracts of the unrelaxed contracting program (2.21) based on Innes (1990).

Lemma 2.4. *Consider the doubly-relaxed contracting program under the bailout regime B. Suppose there is an optimal contract $w^B(y_1)$ that induces $a_1^B > 0$.*

Let μ^B and λ^B denote Lagrange multipliers of (RFOC) and (PC-S), respectively. Thus, $\mu^B \geq 0$. If $\mu^B = 0$, then $a_1^B = a_1^*$. Instead, if $\mu^B > 0$, then (i) $a_1^B < a_1^*$, (ii) for c^B such that $\frac{f_a(c^B|a_1^B)}{f(c^B|a_1^B)} = \frac{\lambda^B - 1}{\mu^B} > 0$, $w^B(y_1)$ has the following form

$$w^B(y_1) = \mathit{1}_{\{y_1 \geq c^B\}} v(y_1) + \mathit{1}_{\{y_1 < c^B\}} A$$

and (iii) $\{a_1^B, w^B(\cdot)\}$ solves the unrelaxed contracting program (2.21).

Proof: See Proposition 2, 3, and 4 of Innes (1990).

Lemma 2.4 states that a live-or-die contract—which pays everything ($v(y_1)$) if the target c^B is met but only the required minimum (the date-2 agency rent A) otherwise—is maximizing the manager's payoff under the bailout regime B if contractual frictions are so large that the incentive constraint is binding ($\mu^B > 0$).

The following theorem shows that if $\mu^B > 0$, then the live-or-die contract is also maximizing incentives. That is, the manager's payoff maximization is equivalent to incentive maximization. Then, *laissez-faire* is best at resolving moral hazard, as any pay restriction can potentially endanger the feasibility of the live-or-die contract.

Theorem 2.2. *Let a_1^B and a_1^{BP} denote equilibrium efforts under the bailout regimes B and BP respectively, where BP uses an arbitrary pay restriction (P) that is not inconsistent with (IC2). Suppose that the doubly-relaxed contracting program under B has a solution $\mathcal{C}^B \equiv \{a_1^B > 0, w^B(y_1)\}$ and $\mu^B > 0$. Then, $a_1^{BP} \leq a_1^B \leq a_1^* < a_1^{**}$*

Proof: Consider the doubly-relaxed contracting program under B. Let λ^B, μ^B, $\theta_L^B(y_1)$, and $\theta_H^B(y_1)$ be the optimal values of Lagrange multipliers of (PC-S), (RFOC), and the two inequality conditions of the boundary constraints (BC), respectively. I can ignore (PC-M) since that $\mathcal{C}^B$ is a solution implies (PC-M). The Kuhn-Tucker conditions are:

$$\lambda^B \left(\int v(y_1) f_a(y_1|a_1^B) dy_1 - c'(a_1^B) \right) + \mu^B \left(\int w^B(y_1) f_{aa}(y_1|a_1^B) dy_1 - c''(a_1^B) \right) = 0 \tag{2.22}$$

$$f(y_1|a_1^B) \left[\mu^B \frac{f_a(y_1|a_1^B)}{f(y_1|a_1^B)} + 1 - \lambda^B \right] = \theta_H^B(y_1) - \theta_L^B(y_1) \quad \forall y_1 \tag{2.23}$$

Next, consider the following exact incentive maximization program:

$$\max_{w:Y\to\mathbb{R}} \ a_1 \quad \text{subject to (PC-M), (PC-S), (BC), and (IC1)}$$

I still use the relaxed first order approach and later verify its validity. For the moment, I ignore (PC-M). Let λ, μ, $\theta_L(y_1)$, and $\theta_H(y_1)$ be Lagrange multipliers of (PC-S), (RFOC), and the two inequality conditions of the boundary constraints (BC), respectively. Lagrange equation is given by

$$\mathcal{L} = a_1 + \lambda \left(\int [v(y_1) - w(y_1)] f(y_1|a_1) dy_1 \right) + \mu \left(\int w(y_1) f_a(y_1|a_1) dy_1 - c'(a_1) \right)$$
$$+ \int \theta_L(y_1) \left[w(y_1) - A \right] dy_1 + \int \theta_H(y_1) [v(y_1) - w(y_1)] dy_1 \tag{2.24}$$

Then, the Kuhn-Tucker conditions are given by

$$\lambda \left(\int v(y_1) f_a(y_1|a_1) dy_1 - c'(a_1) \right) + \mu \left(\int w(y_1) f_{aa}(y_1|a_1) dy_1 - c''(a) \right) = -1 \tag{2.25}$$

$$f(y_1|a_1) \left[\mu \frac{f_a(y_1|a_1)}{f(y_1|a_1)} - \lambda \right] = \theta_H(y_1) - \theta_L(y_1) \quad \forall y_1 \tag{2.26}$$

Let α be such that $\alpha \equiv \left(\int v(y_1) f_a(y_1|a_1^B) dy_1 - c'(a_1^B) \right)^{-1}$, which is positive since $\mu^B > 0$ implies $a_1^B < a_1^*$. Let λ^{BP}, μ^{BP}, $\theta_L^{BP}(y_1)$, and $\theta_H^{BP}(y_1)$ denote $\alpha(\lambda^B - 1)$, $\alpha\mu^B$, $\alpha\theta_L^B(y_1)$, and $\alpha\theta_H^B(y_1)$, respectively. One can check that (i)

$\{\mathcal{C}^B, \lambda^{BP}, \mu^{BP}, \theta_L^{BP}(y_1), \theta_H^{BP}(y_1)\}$ satisfies (2.25) and (2.26). Note that (ii) $\mathcal{C}^B$ satisfies (PC-M) by hypothesis and that (iii) $w^B(y_1)$ is nontrivially increasing. Thus, (i), (ii), and (iii) imply that $\mathcal{C}^B$ is a solution to the exact incentive maximization program.

Note that a_1^{BP} satisfies the four constraints of the exact incentive maximization program and a pay restriction (P). Then, I have $a_1^{BP} \leq a_1^B$ since a_1^B is the highest effort that can be induced from contracts that satisfy the four constraints. Finally, (2.8) implies the desired result. ■

2.3.2 Pay Restrictions Inconsistent with (IC2)

Hitherto, I confine my focus to pay restrictions that are not inconsistent with the date-2 incentive constraint (IC2). If I also consider such pay restrictions, however, the date-1 moral hazard can be mitigated. But it is possible only at the following costs. First, the date-2 moral hazard arises, that is, the date-1 moral hazard is mitigated by exacerbating the date-2 moral hazard. Second, bailouts require more relief fund more often. And third, the quality of shareholder-manager relationship decreases. In the following, I illustrate these points.

Consider the following date-2 pay restriction:

$$|w_2(y_1, g) - w_2(y_1, m)| \leq \Sigma, \quad \Sigma \geq 0, \quad \forall y_1 < D^{BP} \tag{2.27}$$

where Σ is a bonus cap and D^{BP} is a yet to be determined new financial break-even point. Suppose $\Sigma < A \equiv p_h \frac{B}{\Delta p}$. Then, (IC2) can never be satisfied, and hence, the manager will choose 'low' effort at date-2. The ex-post firm value then equals $y_1 + \mathbb{E}[y_2|h] - D$ if $y_1 \geq D^{BP}$ and $y_1 + \mathbb{E}[y_2|l] - D$ otherwise. Thus, (2.27) makes a kink in the ex-post firm value at the critical point D^{BP}.

This means that the downside value is decreased while the upside value is unchanged. The shareholder's willingness to motivate the management thus increases since the shareholder is more willing to avoid to be in the downside. Therefore, one could expect that the date-1 moral hazard is mitigated at the following costs:

(1) Date-2 moral hazard: The manager loses date-2 incentives due to the excessive punishment (2.27).

(2) More often bailouts with larger relief fund: Until now, I assume that the firm cannot borrow at all from the capital market. Suppose now that the firm can raise debt up to α-fraction of its future income $\mathbb{E}[y_2|a_2]$. Thus, this debt capacity increases in the future prospect. The managerial punishment (2.27) decreases this future prospect since the manager could lose the date-2 incentive. The firm then needs to borrow more. To see this formally, note that the firm's new financial break-even point D^{BP} is then given by

$$D^{BP} + \alpha \left\{ \mathit{1}_{\{y_1 \geq D^{BP}\}} \mathbb{E}[y_2|h] + \mathit{1}_{\{y_1 < D^{BP}\}} \mathbb{E}[y_2|l] \right\} = D \qquad (2.28)$$

Recall that the financial break-even point D^B without punishments is determined via

$$D^B + \alpha \mathbb{E}[y_2|h] = D \qquad (2.29)$$

It follows then $D^{BP} > D^B$, as (2.28) and (2.29) imply that

$$D^{BP} - D^B = \alpha \mathit{1}_{\{y_1 < D^{BP}\}} \{\mathbb{E}[y_2|h] - \mathbb{E}[y_2|l]\} > 0$$

If $D^B < y_1 < D^{BP}$, then economically viable firm fails simply because the stringent punishment (2.27). That is, the firm fails more likely under BP than

B, and hence, the government has to undertake bailouts more often. In addition, the government has to lend more relief fund for each $y_1 < D^B$. Under B, it lends $D - \alpha\mathbb{E}[y_2|h] - y_1$. Under BP, however, it has to lend $D - \alpha\mathbb{E}[y_2|l] - y_1$, which is larger than $D - \alpha\mathbb{E}[y_2|h] - y_1$.

(3) The value of shareholder-manager relationship decreases: The manager punishment (2.27) reduces the quality of the shareholder-manager relationship. To see this, consider the following date-0 firm value expression:

$$V(y') = \int_0^{y'} v_l(y_1) f(y_1|a_1(y'))dy_1 + \int_{y'}^{\infty} v_h(y_1) f(y_1|a_1(y'))dy_1 - c(a_1(y'))$$

where $v_{a_2}(y_1) \equiv y_1 + \mathbb{E}[y_2|a_2] - D$ and $a_1(y')$ is the optimal effort that maximizes $V(y')$. For expositional simplicity, suppose that $\bar{\pi}$ equals zero. The equilibrium date-0 firm value then equals $V(y' = 0)$ under B and $V(y' = D^{BP})$ under BP, respectively. By the envelope theorem, it follows that

$$\frac{dV(y')}{dy'} = -[v_h(y') - v_l(y')]f(y'|a_1(y')) = -[\Delta p(g - m)]f(y'|a_1(y')) < 0 \quad \forall y' \tag{2.30}$$

That is, the manager punishment (2.27) reduces the equilibrium date-0 firm value. The reduction in firm value is the larger, the more profitable the date-2 behaving (that is, the more $\Delta p(g - m)$). If the reduction in firm value is substantial, the shareholder-manager relationship is not sustainable or the shareholder has to find a less competitive alternative manager.

2.3.3 Managerial Replacement

Hitherto I consider (any) restrictions on managerial compensation. Another typically used punishment measure is managerial replacement. The bailout literature also considers it frequently. In particular, Bernardo et al. (2011) explicitly examine the effect of managerial replacement on the time-inconsistency problem of government bailouts and show that it mitigates the moral hazard as expected. In this section, I partially counter their result by showing that managerial replacement exacerbates moral hazard if a fired manager's job search cost or the date-2 agency rent is low. Intuitively, there are two related effects of managerial replacement. The first is the usual distortion effect. Managerial replacement is often accompanied with inhibition of severance pay or golden parachute for fired managers, which means shareholders' ability to design long-term compensation scheme is restricted. As I analyze in detail in the previous sections, this distortion effect reduces incentives (especially in the strong governance case). The second is a new effect. If an incumbent manager is fired, he has to find a new job (or begin a start-up). If the job search cost is large, firing incurs large disutility to the manager. This is a direct punishment effect, which increases incentives. The distortion effect dominates the direct punishment effect if the job search cost is little. In the following, this reasoning is formalized.

For simplicity, I consider the following simple version of the general model in Section 2.2; the date-1 output y_1 is binary, that is, $Y = \{0, 1\}$; let a_1 be the success probability $\mathbb{P}(y_1 = 1|a_1)$ and $1 - a_1$ be the failure probability; I assume that the government rescues the firm if $y_1 = 0$; the cost of effort $c(a_1)$ equals $\frac{\gamma}{2}a_1^2$, $\gamma > 1$. Let BR denote the bailout regime conditional on managerial replacement. In the following, I solve for closed form solutions to the contracting programs under B and BR, respectively, and compare the solutions.

Suppose the government chooses BR. If the firm is insolvent at date-1

($y_1 = 0$), the incumbent manager is replaced by a new manager. The new manager has to bear some adjustment cost $\psi \geq 0$ in order to settle in the firm. Let $\{w_2^N(y_2)\}$ denote a contract for the new manager. Then, he will behave if and only if,

$$\mathbb{E}[w_2^N(y_2)|h] - \psi \geq \mathbb{E}[w_2^N(y_2)|l] + B \quad \Leftrightarrow \quad w_2^N(g) - w_2^N(m) \geq \frac{B+\psi}{\Delta p} \quad \text{(IC-N)}$$

I assume for simplicity the shareholder and the new manager' date-2 reservation payoffs are both zero. Facing competition in the labor market for new managers, the shareholder maximizes the new manager's payoff subject to limited liability by paying him everything available:

$$w_2^N(y_2) = y_2 - D, \quad y_2 \in \{g, m\} \tag{2.31}$$

Thus, the date-2 monotonicity constraint is redundant, as the shareholder's ex-post payoff is zero for any y_2. By (2.31), the new manager's incentive constraint (IC-N) is satisfied if and only if

$$g - m \geq \frac{B+\psi}{\Delta p} \tag{2.32}$$

If (2.32) holds, the new manager's equilibrium payoff equals $\mathbb{E}[y_2|h] - D - \psi > 0$. Thus, the new relationship is good (a_2=behaving) if the adjustment cost ψ is moderate. If ψ is substantial, by contrast, the relationship becomes either poor (a_2=misbehaving) or even unsustainable.

At date-1, if $y_1 = 0$, the incumbent manager is fired. He then may find a new job or resort to self-production. Let $\bar{u}_2 \geq 0$ denote the fired manager's date-2 payoff, which is a key parameter in this extension. If the cost of searching a new job is high or the value of self-production is low, then

$\bar{u}_2$ is low. Let u_2 be the manager's equilibrium payoff what he could have obtained if retained ($y_1 = 1$). If $u_2 > \bar{u}_2$, then firing is a punishment. In order to avoid this punishment, the manager will exert a high effort at date-1. This incentive effect of firing will decrease in $\bar{u}_2$.

In the following analysis, it turns out that $\bar{u}_2$ and the agency rent $A \equiv p_h \frac{B}{\Delta p}$ are two key parameters. I confine my focus to the case in which $v(0)$ is the maximum possible level of these parameters. This assumption might be reasonable since $v(0)$ is a large firm's value (at date-2) while $\bar{u}_2$ and A are a small individual's payoff (at date-2). However, this assumption is not crucial for the main result.

At the beginning of date-1, the shareholder offers the incumbent manager a long-term contract $\{w_1^I(y_1), w_2^I(y_1, y_2)\}$, where $w_1^I(y_1)$ is the date-1 compensation given y_1 and $w_2^I(y_1, y_2)$ is the date-2 compensation given y_1 and y_2. Note that $w_1^I(0) = 0$, as the firm has nothing to pay if $y_1 = 0$. I also assume for simplicity that the government prohibits severance pay or the golden parachute, that is, $w_2^I(0, g) = w_2^I(0, m) = 0$. Let $w^I(1)$ denote the long-term compensation $w_1^I(1) + \mathbb{E}[w_2^I(1, y_2)|h]$. Recall that $v(y_1) \equiv y_1 + \mathbb{E}[y_2|h] - D$. The shareholder's participation constraint is equal to

$$\pi(w^I(1)) \equiv a_1[v(1) - w^I(1)] + (1 - a_1)[\underbrace{v(0) - \mathbb{E}[w_2^N(y_2)|h]}_{=0 \text{ by } (2.31)}] = a_1[v(1) - w^I(1)] \geq \bar{\pi} \tag{2.33}$$

and hence the monotonicity constraint is given by

$$w^I(1) \leq v(1) \tag{2.34}$$

The managerial replacement has several important changes on other contractual restraints. First, the incumbent manager's date-1 incentives are not determined by compensation differential but the loss from firing. To see this, given the incumbent manager's long-term payoff

$$u(w^I(1), \bar{u}_2) \equiv a_1 w^I(1) + (1 - a_1)\bar{u}_2 - \frac{\gamma}{2}a^2 \tag{2.35}$$

the date-1 incentive constraint is given by

$$\gamma a_1 = [w^I(1) - \bar{u}_2] \tag{IC-I}$$

where I can interpret $[w^I(1) - \bar{u}_2]$ as the loss from firing.

Second, the adverse effect of the date-2 agency rent A on the date-1 incentives can be mitigated. To see this, recall that the limited liability restrictions and the incumbent manager's date-2 incentive constraint end up restricting $w^I(1)$ in the following way:

$$A \leq w^I(1) \leq v(1) \tag{2.36}$$

Note that the above restriction on long-term compensation applies only for the success case $y_1 = 1$. In contrast, under B (without replacement), the incumbent manager must be paid at least A for any $y_1 \in \{0, 1\}$. Since such an insurance for the incumbent manager is disappeared under BR (with replacement), he faces greater incentives *ceteris paribus*.

I assume for simplicity the incumbent manager's date-1 reservation payoff is low enough that his participation constraint is redundant. Then, the long-term contracting program under BR at date-1 is given by

$$\max_{w^I(1)} \; u \quad \text{subject to (2.33), (2.34), (2.35), (IC-I), and (2.36)}$$

The existence of a solution to the BR contracting program depends on $\bar{\pi}$ and A in a rather complicated way. See the attached footnote for the detail.[20] For expositional simplicity, I assume that $p_h(g-m) < 1$[21] and $m = D$. The following proposition provides some essential feature of the solutions to the contracting programs under the bailout regimes B and BR, respectively.

Proposition 2.1. *Suppose $p_h(g-m) < 1$ and $m = D$ for expositional simplicity.*

(i) Under BR, the shareholder-manager relationship is sustainable if

$$\bar{\pi} \leq \frac{[v(1) - \bar{u}_2]^2}{4\gamma}$$

If the relationship is sustainable, the equilibrium effort $a_1^{BR}(\bar{\pi})$ is given by

$$a_1^{BR}(\bar{\pi}) = \frac{(v(1) - \bar{u}_2) + \sqrt{(v(1) - \bar{u}_2)^2 - 4\bar{\pi}\gamma}}{2\gamma}$$

(ii) Under B, the shareholder-manager relationship is sustainable if

$$\bar{\pi} \leq v(0) - A + \frac{1}{4\gamma}$$

If the relationship is sustainable, the equilibrium effort $a_1^B(\bar{\pi})$ is given by

$$a_1^B(\bar{\pi}) = \mathit{1}_{\{\bar{\pi} \leq v(0)-A\}} \frac{v(1) - v(0)}{\gamma} + \mathit{1}_{\{\bar{\pi} > v(0)-A\}} \frac{v(1) - v(0) + \sqrt{v(1) - v(0) - 4\left(\bar{\pi} - \{v(0) - A\}\right)}}{2\gamma}$$

---

[20] There are two cases to consider. First, if $A \leq \frac{v(1)+\bar{u}_2}{2}$, a solution exists for any $\bar{\pi} \leq \frac{[v(1)-\bar{u}_2]^2}{4\gamma}$. Second, if $A > \frac{v(1)+\bar{u}_2}{2}$, then a solution exists for any $\bar{\pi} \leq \pi(A) < \frac{[v(1)-\bar{u}_2]^2}{4\gamma}$, where $\pi\left(w^I(1) = A\right)$ is defined in (2.33). To understand this second case, note that as $\bar{\pi}$ increases, the shareholder has to reduce the payment $w^I(1)$ for the incumbent. But $w^I(1)$ cannot be lower than the agency rent A. Thus, the relationship breaks if A and $\bar{\pi}$ are high enough.

[21] If $p_h(g-m) < 1$, the existence condition becomes simpler, as the second case in the footnote above is disappeared: I assume that $v(0) = p_h(g-m) + m - D$ is the maximum possible level of $\bar{u}_2$ and A. Then, the second case is not satisfied if $p_h(g-m) < 1$ and $m = D$

Proof:(i) Let w' denote $w^I(1)$. The BR contracting program can be simplified to:

$$\max_{A \leq w' \leq v(1)} \quad \frac{1}{2\gamma}[w' - \bar{u}_2] + \bar{u}_2 \quad \text{subject to}$$

$$\pi(w') = \frac{1}{\gamma}[w' - \bar{u}_2][v(1) - w'] \geq \bar{\pi}$$

See Figure 2.4. Note that $\pi(w')$ is inverse U-shaped and strictly concave, $\pi(\bar{u}_2) = \pi(v(1)) = 0$ and $\max\{\pi(w')\} = \pi\left(\frac{v(1)+\bar{u}_2}{2}\right) = \frac{(v(1)-\bar{u}_2)^2}{4\gamma}$. If $A \leq \frac{v(1)+\bar{u}_2}{2}$, the solution exists if $\bar{\pi} \leq \frac{(v(1)-\bar{u}_2)^2}{4\gamma}$. If $A > \frac{v(1)+\bar{u}_2}{2}$, the solution exist if $\bar{\pi} \leq \pi(A)$. Since $\bar{u}_2 < \frac{v(1)+\bar{u}_2}{2}$, the solution, if exists, is given by

$$\arg\max\{w' : A \leq w' \leq v(1), \pi(w') = \bar{\pi}\}$$

which is equal to

$$\frac{(v(1) + \bar{u}_2) + \sqrt{(v(1) - \bar{u}_2)^2 - 4\bar{\pi}\gamma}}{2} \in \left[\frac{(v(1) + \bar{u}_2)}{2}, v(1)\right]$$

Then, the equilibrium effort in Proposition 2.1 is given by (IC-I).

(ii) Let w denote $w(0)$ and Δ denote $w(1) - w(0)$. Assume that $\Delta \geq 0$ in equilibrium (and it will be verified.) By substituting Δ for a, the contracting program under B is reduced to

$$\max_{\Delta, w} \frac{\Delta^2}{2\gamma} + w(0) \quad \textit{subject to}$$

$$\frac{\Delta}{\gamma}[1 - \Delta] + v(0) - w(0) \geq \bar{\pi}$$

$$A \leq w(0) \leq v(0), \quad \Delta + w(0) \leq v(1)$$

Lagrange equation L equals

$$L = \frac{\Delta^2}{2\gamma} + w(0) + \lambda\left(\frac{\Delta}{\gamma}[1 - \Delta] + v(0) - w(0) - \bar{\pi}\right)$$

$$+\theta_L(w(0)-A)+\theta_H(v(0)-w(0))+\theta_3(v(1)-\Delta-w(0))$$

where λ, θ_L, θ_H, and θ_3 are all nonnegative relevant multipliers. Kuhn-Tucker conditions are

$$\Delta+\lambda(1-2\Delta)=\theta_3 \tag{2.37}$$

$$1-\lambda+\theta_L-\theta_H-\theta_3=0 \tag{2.38}$$

$$\lambda\left(\frac{\Delta}{\gamma}[1-\Delta]+v(0)-w(0)-\bar{\pi}\right)=0 \tag{2.39}$$

$$\lambda\geq 0,\quad \frac{\Delta}{\gamma}[1-\Delta]+v(0)-w(0)\geq\bar{\pi} \tag{2.40}$$

$$\theta_L(w(0)-A)=0,\quad \theta_H(v(0)-w(0))=0,\quad \theta_3(v(1)-\Delta-w(0))=0 \tag{2.41}$$

$$\theta_L\geq 0,\quad \theta_H\geq 0,\quad \theta_3\geq 0,\quad w(0)\geq A,\quad w(0)\leq v(0),\quad \Delta+w(0)\leq v(1) \tag{2.42}$$

Note that the shareholder's participation is a main driving force that causes contractual frictions. Thus, it should be suboptimal to give rents to the shareholder. In the following, I guess that the shareholder's participation is binding (i.e. $\lambda>0$), and it will be verified.

The total surplus maximizing effort a_1^* equals $\frac{1}{\gamma}$, which can be induced if Δ equals 1. I first find parameter values of $\bar{\pi}$ such that Δ equals 1 in the equilibrium. If $\Delta=1$, (2.39) implies $w(0)=v(0)-\bar{\pi}$ since $\lambda>0$. Then, since $w(0)\geq 0$, it follows that $\bar{\pi}\leq v(0)$. The following $\{\Delta, w(0), \lambda, \theta_L, \theta_H, \theta_3\}$ satisfies all of Kuhn-Tucker conditions and hence constitutes the solution:

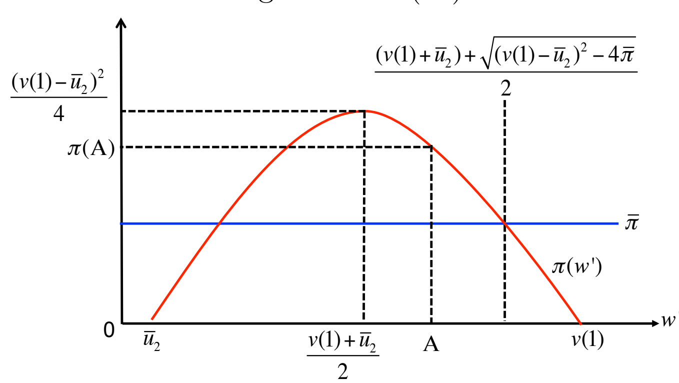
$$\Delta=1,\quad w(0)=v(0)-\bar{\pi},\quad \lambda=1,\quad \theta_L=\theta_H=\theta_3=0,\quad \text{if } 0\leq\bar{\pi}\leq v(0)-A \tag{2.43}$$

Next I consider the case where $\bar{\pi} > v(0) - A$. If $\bar{\pi}$ is very high, it is obvious that the shareholder-manager relationship cannot be sustained. I shall find such a critical level of $\bar{\pi}$. (2.40) implies that the shareholder's payoff is maximized if $(\Delta, w(0)) = \left(\frac{1}{2}, A\right)$, and with this contract, the payoff equals $v(0)+\frac{1}{4\gamma}$. Therefore, the relationship can be sustained only if $\bar{\pi} \leq v(0)-A+\frac{1}{4\gamma}$. Note that if $\bar{\pi} \leq v(0) - A$, the optimal salary $w(0) = v(0) - \bar{\pi}$ is decreasing in $\bar{\pi}$ and hits a_1 at $\bar{\pi} = v(0)-A$. Then, one can guess that it remains at a_1 if $\bar{\pi} > v(0) - A$. Given $w(0) = A$ and $\lambda > 0$, (2.39) implies $\Delta = \frac{1+\sqrt{1-4\gamma(\bar{\pi}-\{v(0)-A\})}}{2}$. The following $\{\Delta, w(0), \lambda, \theta_L, \theta_H, \theta_3\}$ satisfies all of Kuhn-Tucker conditions and thus constitutes the solution:

$$\Delta = \frac{1+\sqrt{1-4\gamma(\bar{\pi}-\{v(0)-A\})}}{2} \in \left(\frac{1}{2}, 1\right), \quad w(0) = A, \quad \lambda = \frac{\Delta}{2\Delta - 1} > 0,$$
$$\theta_L = \frac{1-\Delta}{2\Delta-1} > 0, \quad \theta_H = \theta_3 = 0, \quad \text{if } v(0) - A < \bar{\pi} < v(0) - A + \frac{1}{4\gamma} \tag{2.44}$$

Since $a_1 = \Delta$, (2.43) and (2.44) imply the desired result. ■

Figure 2.4: $\pi(w')$



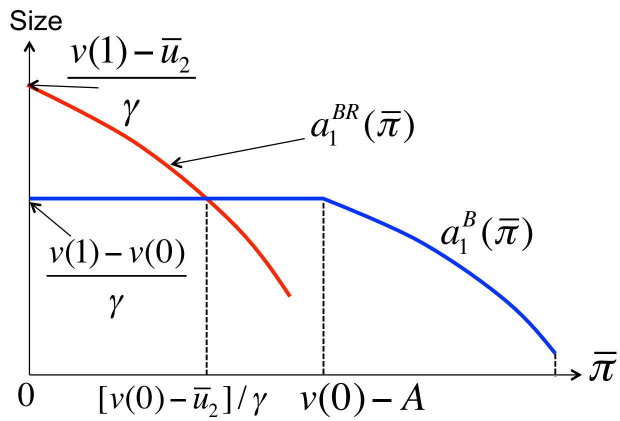
Proposition 2.1 imply that the relative efficiency of the bailout regime B over the bailout regime BR depends on two key parameters: the incumbent

manager's date-2 payoff after firing $\bar{u}_2$ and the expected agency rent A. B induces the higher effort, the smaller the agency rent A, since then the insurance for failed manager is smaller. Recall that under BR the agency rent does not affect incentives since it is no more an insurance for the incumbent manager: if he fails, he will be fired and hence he cannot guarantee A. BR induces the higher effort, the smaller $\bar{u}_2$, since then the loss from firing is the larger. That is, $\frac{\bar{u}_2}{A}$ is a crude measure of the relative efficiency of B in comparison to BR.

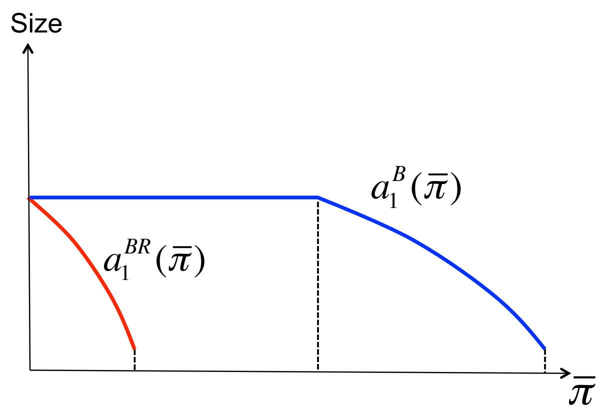
Recall that $v(0)$ is the maximum possible level of $\bar{u}_2$ and A. Then, the following four extreme cases (and related figures) illustrate when managerial replacement exacerbates moral hazard. I first consider the two cases in which there is no agency rent, that is, $A = 0$.

(a) $\bar{u}_2 = 0$: First, see $a_1^B(\bar{\pi})$ in Figure 2.5(a). As bailouts resolve the liquidity shortage problem, the firm gets the bailout rent $v(0)$. If there are no contractual frictions ($\bar{\pi} = 0$), it is optimal to maximize total surplus and let the manager take it all. The equilibrium effort than equals the total surplus maximizing effort. As the (ex-post) firm value equals $v(1)$ and $v(0)$ on the upside and the downside, respectively, the equilibrium effort equals $[v(1) - v(0)]$. If $\bar{\pi} > 0$, then the manager has to transfer a fraction of surplus (with the size $\bar{\pi}$) to the manager. There are two ways to transfer: the base pay $w(0)$ cut and the incentive pay $(w(1) - w(0))$ cut. Note that the bailout rent $v(0)$ is a buffer for the base pay: the limited liability restriction implies that $w(0) \leq v(0)$. Then, an optimal way to transfer surplus is to cut the base pay but not the incentive pay. If $\bar{\pi}$ is sufficiently large, then the base pay hits zero and hence the incentive pay should also be cut. $a_1^B(\bar{\pi})$ reflects this reasoning: a_1^B is constant as $v(1) - v(0)$ for some moderate levels of $\bar{\pi}$ and then gradually

Figure 2.5: $a_1^{BR}(\bar{\pi})$ vs. $a_1^{B}(\bar{\pi})$



(a) $\bar{u}_2 = 0$ and $A = 0$

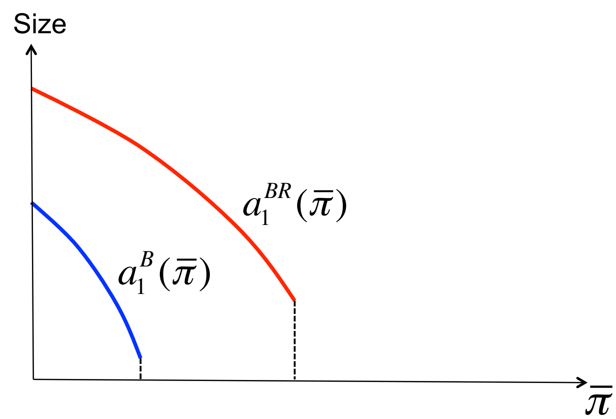


(b) $\bar{u}_2 = v(0)$ and $A = 0$

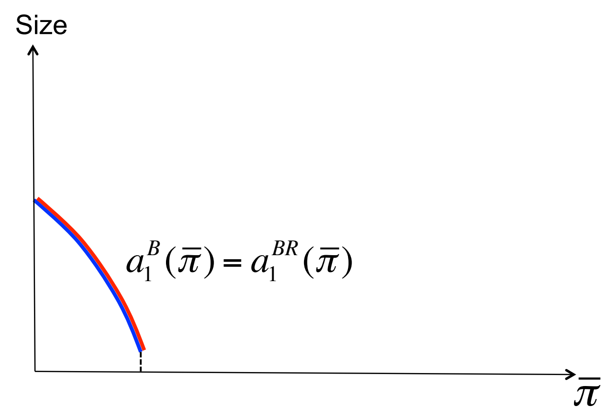
decreases in $\bar{\pi}$.

Next, see $a_1^{BR}(\bar{\pi})$ in Figure 2.5(a). Managerial replacement essentially reduces the manager's downside wage from $w(0)$ to $\bar{u}_2 = 0$. The (ex-post) firm value generated by the shareholder and the incumbent manager is $v(1)$ on the upside and $\bar{u}_2 = 0$ on the downside. In the frictionless case of $\bar{\pi} = 0$, thus $v(1)$ is the equilibrium incentive, which is greater than the equilibrium

Figure 2.6: $a_1^{BR}(\bar{\pi})$ vs. $a_1^B(\bar{\pi})$



(a) $\bar{u}_2 = 0$ and $A = v(0)$



(b) $\bar{u}_2 = v(0)$ and $A = v(0)$

incentive under B. If contractual frictions exist, the manager has to transfer $\bar{\pi}$ to the shareholder. Since the wage on the downside is fixed as $\bar{u}$, the only way to transfer surplus is by cutting the incentive pay $w^I(1) - \bar{u} = w^I(1)$. Thus, equilibrium incentive decreases strictly in $\bar{\pi}$ whenever $\bar{\pi} > 0$. In consequence, managerial replacement ends up mitigating moral hazard if contractual frictions are moderate, whereas it ends up exacerbating moral hazard otherwise.

(b) $\bar{u}_2 = v(0)$: As $\bar{u}_2$ jumps up to $v(0)$, the total surplus maximizing effort decreases to $v(1) - \bar{u}_2 = v(1) - v(0)$. Thus, the equilibrium incentive in the frictionless case is equal to $v(1) - v(0)$. The equilibrium incentive is strictly decreasing in $\bar{\pi}$ since the only way to transfer surplus is by cutting the incentive pay $w^I(0) - \bar{u}_2$. In this case, managerial replacement exacerbates moral hazard for any $\bar{\pi} > 0$.

Next, I consider two extreme cases in which the agency rent A is at its maximum: $A = v(0)$.

(c) $\bar{u}_2 = 0$: Note that the agency rent A is the lower bound for the long term wage $w(y_1)$ under B. Thus, the base pay $w(0)$ cannot be less than $A = v(0)$. Under B, thus, the base pay must be $v(0)$ for any $\bar{\pi}$, and hence, the incentive pay must be reduced as $\bar{\pi}$ increases.

(d) $\bar{u}_2 = v(0)$: In this case, both regimes induce their lowest efforts. Interestingly, they induce the same effort for any $\bar{\pi}$.

In sum, managerial replacement could exacerbate moral hazard if the job search cost is low, the agency rent is low, or the contractual frictions are high.

2.3.4 Subjective Performance Evaluation

Until now, I focus on managerial compensation and assume that output y_t is a single natural performance measure. However, in the real world, there is the problem of measurement error both for management-level and employee-level compensations. There are various performance measures, such as earnings, stock price, sales, contribution to consumer satisfaction, relationship with customers, leadership, etc. Some measures are objective and hence verifiable by third-parties like courts or governments, whereas others are subjective and thereby nonverifiable. Unfortunately, no measures can exactly capture the true performance such as contribution to long-term firm value. This is particularly the case if only objective measures are used in evaluation.

Murphy (2009) observes a related anecdotal event: before the financial crisis, Washington Mutual rewarded brokers if they sell mortgage loans. Importantly, bonuses were based on the quantity but not the quality of loan sales. Consequently, employees sold a large amount of loans without exerting much effort in assessing debtors' financial ability of repayment, and finally, the company collapsed. Countrywide Finance and Wachovia and other smaller lenders experienced similar events. An implication from this anecdote is that there is a disparity between objective performance measure (such as sales of loan) and true performance measure (such as sales of 'good' loan). To reduce the disparity, firms often need to check various sources of subjective information such as whether a worker is sincere and not manipulating objective measures by letting the firm value at risk.

Along this line, the contract theory literature (especially relational contract theory) emphasizes the importance of subjective performance evaluation. Relatedly, the informativeness principle provided by Hwang (2014) implies that optimal compensation should be based on a balanced score card, which is a collection of internal performance measures. However, government bailout policies (such as United States Troubled Asset Relief Program) demand transparency since they lost confidence on internal-firm decision-making and imposed objective-measure dependent pay restrictions. Firms then cannot utilize valuable but subjective internal information, which is helpful in reducing the measurement problem. Moral hazard is then exacerbated.

2.3.5 Bargaining Power and Risk Aversion

Bargaining power: Throughout this paper, I assume that the labor market is competitive, and hence, management has full bargaining power. Suppose instead that shareholders have full bargaining power: they maximize their own payoff rather than the manager's one. Constraints to the contract design problem are the same but shareholders' participation constraint is replaced by manager's participation constraint. Thus, the manager's reservation payoff $\bar{u}$ is a key parameter. (I assume $\bar{\pi} = 0$ so that shareholders' participation is automatically satisfied). In this case, augmented stock options are still optimal in the strong governance case. Any restrictions on managerial compensation then hamper the availability of augmented stock options, and hence, potentially reduce the date-1 incentives.

Risk-averse manager: Suppose the manager is risk-averse and there are no limited liability restrictions. Managerial punishments still play new constraints to the contract design problem. The total surplus of shareholders and manager then decreases. Thus, the shareholder-manager relationship becomes unsustainable and hence the manager leaves if their reservation payoffs are sufficiently high. In response, the shareholders will hire inferior alternative management. One might be more interested in the case where the reservation payoffs are so low that the relationship still sustains even with managerial punishments. Relatedly, Jewitt et al. (2008) examine how boundary conditions affect optimal contracting in the risk-aversion model and conclude that it is a real challenge to understand how contractual environment (such as minimum pay restrictions) affects equilibrium effort when the agent (manager) is risk-averse.

2.4 Shareholder or Debtholder Punishments

Managerial punishments (or regulations on pay structure) could exacerbate moral hazard because they distort shareholders' ability to motivate management though incentive design (as they are new constraints to the contracting problem) while they have no effect in shareholders' willingness to motivate (as shareholders are not directly punished).

Natural alternatives are shareholder punishments. If shareholders expect punishments when their firm fails, then their ex-ante willingness to motivate management increases since they need to avoid the punishments. In my model, shareholders' willingness to motivate is captured by the slope of firm value $v(y_1)$ or the difference between upside and downside firm values. Shareholder punishments reduce the firm value on the downside (i.e. $\{y_1 : y_1 < D\}$), whereas the firm value on the upside is unaffected. Accordingly, their willingness to motivate management increases. However, their ability to motivate is unaffected, as shareholder punishments do not restrict the functional form of long-term contracts. Therefore, shareholders will choose more powerful contracts and hence moral hazard is mitigated.

Similarly, debtholder punishments can be effective in mitigating moral hazard. If debtholders expect punishments when their debtor-firm fails, then their ex-ante monitoring incentives increase. They may set more stringent covenants that secure the firm's sustainability. A standard result in contract theory is that monitoring reduces contractual inefficiency (see Holmstrom (1979), Kim (1995) and also Hwang (2014)). The current model can easily be extended to accommodate this monitoring feature. Let $\alpha_1 \in \{l, h\}$, $h > l \geq 0$, denote the monitoring effort of debtholders. Suppose that if debtholders' monitoring effort α_1 increases, the information structure $F(y_1|a_1, \alpha_1)$ becomes more informative in the sense of Fisher information, that is,

$$\text{Var}\left(\frac{f_a(y_1|a_1, h)}{f(y_1|a_1, h)}\right) > \text{Var}\left(\frac{f_a(y_1|a_1, l)}{f(y_1|a_1, l)}\right)$$

where Var is variance. Hwang (2014) shows that in this case the equilibrium effort increases in debtholders' monitoring effort.

There are various real world examples of shareholder or debtholder punishments. During the recent 2008-09 financial crisis, debt-for-equity swap attracts a lot of attention.[22] The swap is implemented by using contingent convertible bonds (CoCos). CoCo is a convertible bond, which is initially a debt contract. If the firm is financially distressed (so that predetermined trigger conditions are satisfied), the debt is swapped for equity. This is effective punishment for both shareholders and debtholders. For shareholders, their existing equity value is diluted if the firm is in distress. For debtholders, they have to suffer from unwanted swap since the equity value of a distressed firm is low. Therefore, moral hazard is expected to be relieved.

However, CoCos are involved with measurement errors. Debt is swapped for equity if predetermined trigger conditions are satisfied. Trigger conditions are often based on accounting profit or stock price. Accounting profit is easily manipulable. Stock price is not a good measure of firm's financial distress or likelihood of bankruptcy. Instead, Oliver Hart and Luigi Zingales propose to use the credit default swap (CDS) spread in determining trigger conditions.[23]

[22]See Pazarbasioglu et al. (2011).
[23]See "How to make a bank raise equity," Financial Times, Feb. 7, 2010.

CDS is an insurance contract, which pays debtholders principal and interests when their debtor-firm is in default. Thus, the market spread of CDS is a good measure of a firm's financial difficulty. Also, it is more difficult to manipulate than accounting profit.

Capital reduction is a traditional measure of shareholder punishments. As a condition of bailouts, governments may require shareholders to reduce equity. Acharya and Yorulmazer (2008) advocate partial capital reduction over total capital reduction since total capital reduction replaces management and board of directors altogether, and hence, their knowledge or know-how can be lost.

A natural critique against these shareholder or debtholder punishments is that they increase firms' cost of capital: if investors suffer from some punishment measures, they are reluctant to invest. This will be certainly the case in partial equilibrium analysis. In general equilibrium analysis, however, cost of capital does not necessarily increase: these punishments (as prudential regulations) reduce the systemic risk of financial system. Investors thus require lower required rate of return, and hence, firms' cost of capital may decrease.

2.5 Conclusion

After the recent 2008-09 financial crisis, the public, regulators, market participants and political leaders have increasingly called for direct government regulation on managerial (or employee) compensation. This call is aimed at mitigating moral hazard (such as excessive risk-taking or imprudent risk-management), which is the seed of next bailouts, or at curbing excessively high pay for top management.

However, this paper shows that direct regulation on compensation scheme could cause unforeseen adverse effects such as more severe moral hazard. This is because pay restrictions could distort the incentive contracting problem: shareholders' ability to motivate management is reduced since pay restrictions are new constraints to the contract design problem, whereas their willingness to do so is unaffected since shareholders are not directly punished. Also, objective-measure based pay restrictions disallow the use of valuable but subjective internal information in motivating employees.

As alternatives, I advocate punishments or prudential regulation of shareholders such as debt-for-equity swap, CDS margin call and/or partial equity reduction. These measures are effective in mitigating moral hazard: shareholders' willingness to motivate management increases since they are punished, whereas their ability to motivate is not affected since shareholder punishments are not constraints to the contract design problem. Similarly, debtholder punishments increase debtholders' incentives to monitor shareholders and management so that they design contracts in a way to reduce the default risk. This suggestion would be more appealing if one acknowledges that outside parties like regulators can hardly observe the detail of complicated within-firm decision-making. That is, shareholders (or debtholders) should be given enough freedom to make internal decision but they should be responsible for the outcome.

Chapter 3

Education and Career-Concerns

3.1 Introduction

Life consists of two important stages. In the first stage, as students, people decide how much education to take. In the second stage, as workers, they choose jobs and exert efforts in workplace. (And then they retire). Therefore, education and career are two important choices people make in life. But people do not choose education and career independently. They choose education anticipating its impact on future careers. However, the existing literature rarely considers possible interactions between education and career. A standard model of (stand-alone) education is Spence (1973) job market signaling model in which education is used to signal workers' hidden productivity to the labor market. A standard model of (stand-alone) career is Holmstrom (1999) career-concerns model in which workers exert efforts to make good performance, which appeals to the labor market that they are talented workers. Although the education literature and the career-concerns literature starting from the two seminal papers each are both vast, little work has been done on the interactions of education and career-concerns.

This paper examines an integrated model of education and career-concerns. By explicitly considering interactions between education and career-concerns, I find two new results on education. First, if the society uses education to sort highly productive agents from lowly productive agents, then education could reduce the total surplus. This is inconsistent with Spence (1973)'s standard job market signaling model, which implies that education as a sorting device is welfare-enhancing (even if it has no human capital value). Suppose there are high-type and low-type agents. The labor market treats them equally, as they are indistinguishable. Then, high-type leaves and only low-type prevails in the labor market (with appropriate assumption on reservation payoffs). This loss of top talent (i.e. the lemon market problem) can be resolved if education reveals hidden types.

This story, however, may be overturned if not just hidden productivity (type) but also hidden effort contributes to output. Suppose in the post-education work-stage, an agent exerts effort to produce output. If the labor market does not fully know the agent's hidden productivity, she works hard to demonstrate her talent. But if education reveals the hidden talent, she is demotivated, exerts little effort, and hence, output decreases in expectation. Therefore, education as a sorting device could be detrimental to welfare. This result implies that governmental education support must be universal rather than selective. Suppose that education improves workers' productivity. Governmental education support such as nationwide scholarship programs then contribute to the total surplus by increasing overall human capital value. Government can implement it as an universal measure, which applies to the best majority of students without much preconditions, or as a selective measure, which helps only those with high grades, scores, or other indicators that demonstrate their potentials. With selective measures, education is used as a human-capital enhancing device, which is beneficial to social welfare, on the one hand, and as a sorting device, which is detrimental to social welfare, on the other hand. Thus, education support policies should be used with universal measures.

The second result shows that education can be used to hedge income-risk. I begin with a related anecdote of Marissa Mayer, the CEO of Yahoo!. After taking a Master's degree from Stanford University in 1999, she received many job offers including a teaching job at Carnegie Melon University, a consulting job at McKinsey and Company, and an engineer job at a small start-up with only 20 employees. She chose the startup, which has a weird name 'Google,' and made a big success story thereafter. Why some people like Marissa Mayer, Steve Jobs, Mark Zuckerberg, and many other smart people choose high-return but high-risk jobs such as entrepreneurs in start-ups or traders in investment banks? Why some other people choose medium-return but low-risk jobs such as government agencies or think-tanks? How education, careers, and job-risk are connected? The second result addresses these questions. In order to receive high pay in the labor market, people can choose one of two ways. One way is taking high education (i.e. Ph.D. degree) in order to separate themselves from low-ability workers. The labor market then treat them well knowing that they are of high-ability workers. The other way is to take little education (i.e. College or Master's degree), begin career early, and show good workplace performance. This way is risky, as performance depends not only ability but also other exogenous shocks, but it promises great return if performance is good. Thus, those with high risk-aversion choose high-risk but high-return jobs by taking little education, while those with low risk-aversion choose safe but medium-return jobs by taking high education. Thus, education is used to hedge job-risk. This result is related to the large literature of education. The existing literature considers human-capital enhancement and sorting hidden types as two main roles of education. (See Weiss (1995) for related survey). But this paper suggests a third role of education, a job-risk hedging device.

This paper is organized as follows. In section 3.2, I provide an integrated model of education and career-concerns and find the welfare implication of education. In section 3.3, I find the new role of education as a job-risk hedging device.

3.2 Model I: Welfare Effect of Education

An agent's hidden productivity θ is either h (high) or l (low) where $0 = l < h < 1$. There are half-measure of high-type and half-measure of low-type agents.

Timeline is as follows. There are three periods $t = 0, 1, 2$. At $t = 0$, each agent as a student chooses publicly-observable education level $e \in [0, \infty)$ at the cost of education $C(\theta, e)$ such that $C(\theta, 0) = 0$, $C_e > 0$, $C_{e\theta} < 0$ where lower subscript denotes (cross-) partial derivative. At $t = 1, 2$, competitive labor market pays wage w_t; Given w_t, each agent as a worker chooses privately-observable effort level $a_t \in [0, \bar{a}]$ at the cost of effort $c(a_t)$ such that $c(0) = c'(0) = 0$, $c' > 0$, $c'' > 0$; Given a_t, publicly-observable output $y_t \in \{0, 1\}$ is realized.

Effort and productivity contribute to output in the following manner: let $f(\theta, a_t)$ be the conditional probability of success

$$f(\theta, a_t) \equiv \mathbb{P}(y_t = 1|\theta, a_t) = \theta + ka_t, \;\; k > 0 \tag{3.1}$$

Note that the success probability increases in effort and productivity. (The following analysis holds even if $f(\theta, a_t)$ has a general functional form such that

$f_a > 0$ and $f_\theta > 0$). To ensure that $f(\theta, a_t) \leq 1$ for any $a_t \in [0, \bar{a}]$, I need to assume that $\bar{a} \leq \frac{1-h}{k}$. Since $y_t \in \{0, 1\}$, one can interpret $f(\theta, a_t)$ as the expected output conditional on (θ, a_t).

I assume that labor market is competitive and hence wage w_t equals market's expectation of output. I assume that output y_t is not contractible. A such case is as follows. Explicit incentive contracts result in wage inequality across agents working for a same employer. The employer will be afraid of resulting envy or demoralization.

Given wage structure, a type-θ agent's preference is given by

$$-C(\theta, e) + \mathbb{E}^\theta[w_1|\cdot] - c(a_1) + \delta \left\{\mathbb{E}^\theta[w_2|\cdot] - c(a_2)\right\} \tag{3.2}$$

where $\mathbb{E}^\theta[\cdot|\cdot]$ denotes the type-θ agent's expectation with respect to relevant information structure.

In the following, we consider two cases. First, high-type agents choose higher level of education than low-type agents—separating equilibrium. Second, both types choose the same level of education—pooling equilibrium.

3.2.1 Separating equilibrium

Note that education has no human capital value. It can only signal hidden productivity. Thus, the most reasonable case is such that low-type chooses zero education while high-type chooses the minimum positive education, which low-type cannot mimic and hence hidden types are revealed. This separating equilibrium is called Riley equilibrium. I confine my attention to Riley equilibrium, which is later shown to be the only separating equilibrium that satisfies the Intuitive Criterion suggested by Cho and Kreps (1987).

In Riley equilibrium, types are revealed and so is expected output. Suppose market expects $a_t^e(\theta)$ will be chosen in equilibrium. A type-θ agent's wage is then, for $t = 1, 2$,

$$w_t = f(\theta, a_t^e(\theta)) = \theta + k a_t^e(\theta)$$

That is, wage is fixed and independent of effort a_t (though it depends on $a_t^e(\cdot)$). Let $a_t^*(\theta)$ denote the equilibrium effort type-θ agent actually chooses. Since effort is costly, it follows $a_t^*(\theta) = a_t^e(\theta) = 0$ in rational expectations equilibrium.

Let $e(\theta)$ be the equilibrium level of education, which type-θ agent chooses. Thus, $e(l) = 0$ and $e(h) > 0$ is the minimum education level low-type cannot profitably mimic. That is, $e(h)$ is characterized by

$$h(1+\delta) - C(l, e(h)) = l(1+\delta) - C(l, e(l)) = 0 \tag{3.3}$$

Let $u(\theta)$ denote type-θ agent's reservation payoff, which she can get by leaving the labor market. This is perhaps the payoff of self-production or participating in a secondary labor market. I assume $u^l = 0$ by normalization and that

Assumption 3.1. $\frac{1}{2}h < u^h < h$

Assumption 3.1 implies the lemon market problem: high-type agents will not participate unless types are revealed, though their participation is efficient. I also assume that education is not prohibitively costly:

Assumption 3.2. $(1+\delta)[h - u^h] > C(h, e(h))$

(3.3) implies that Assumption 3.2 holds if education cost decreases sufficiently in productivity.

The total surplus in Riley equilibrium is given by

$$\frac{1}{2}\{(1+\delta)h - C(h, e(h))\} \tag{3.4}$$

As a benchmark case, suppose that education cannot be used as a sorting device and there is no post-education work-stage. By the lemon market problem then the total surplus is given by

$$\frac{1}{2}\left\{(1+\delta)u^h\right\} \tag{3.5}$$

Assumption 3.2, (3.4) and (3.5) imply that education as a sorting device mitigates the lemon market problem by inducing high-type agents' participation in the labor market. Furthermore, education as a sorting device improves efficiency in Pareto sense.

3.2.2 Pooling equilibrium

In a pooling equilibrium, education has no sorting effect and hence the labor market cannot distinguish types. Since education has no role but only costly, I confine my attention to the most efficient pooling equilibrium in which both types choose zero education. In Spence's standard job market signaling model where there is no post-education work-stage, this pooling equilibrium fails to satisfy the Intuitive Criterion. Later, I shall show that it is not the case if we add post-education work-stage.

Suppose market expects that $a_t^e(\theta)$ is chosen in equilibrium. Let $\mathbf{a}_t^e \equiv (a_t^e(h), a_t^e(l))$. Let $\mathbb{E}^m$ denote market's expectation conditional on $\mathbf{a}_t^e$. Then, wages are given by

$$w_1 = \mathbb{E}^m[y_1] \tag{3.6}$$

$$w_2(y_1) = \mathbb{E}^m[y_2|y_1] \tag{3.7}$$

Importantly, date-2 wage (i.e. expected date-2 output) depends on date-1 output y_1. This is because market observes date-1 performance y_1, which is informative of an agent's hidden productivity θ, which determines date-2 performance y_2. (In a separating equilibrium, in contrast, date-1 performance has no value of information, as productivity is already unraveled). This observation is crucial in this paper. Even if there are no explicit incentive contracts, agents face date-1 incentives. This is because date-1 effort (stochastically) determines date-1 output, which determine date-2 wage. However, date-1 wage is fixed and independent of output, as no output is realized in the beginning of date-1.

To see this formally, note that at date-2 a type-θ agent maximizes $w_2(y_1) - c(a_2)$. Since $w_2(y_1)$ is independent of a_2 (while it depends on $a_2^e(\theta)$), the agent chooses $a_2^*(\theta) = 0$ for $\theta = h, l$. Market's rational expectation ends up $a_2^e(\theta) = a_2^*(\theta) = 0$. At date-1, the agent chooses $a_1 = a_1^*(\theta)$, which maximizes

$$w_1 - c(a_1) + \delta \left\{ \mathbb{E}^\theta[w_2(y_1)|a_1] - c(a_2) \right\} \tag{3.8}$$

where $\delta \in (0,1)$ is discount factor and $\mathbb{E}^\theta$ is type-θ agent's expectation conditional on θ and $\mathbf{a}_t^e$. Since a_1 affects the probability of y_1, $a_1 = a_1^*(\theta)$ solves

$$\begin{aligned} c'(a_1) &= \delta \frac{\partial}{\partial a_1} \mathbb{E}^\theta[w_2(y_1)|a_1] \\ &= \delta \frac{\partial}{\partial a_1} \left[f(\theta, a_1) w_2(1) + \{1 - f(\theta, a_1)\} w_2(0) \right] \\ &= \delta f_a(\theta, a_1)[w_2(1) - w_2(0)] : \text{marginal incentive} \end{aligned} \tag{3.9}$$

In general, $a_1^*(\theta)$ depends on θ, as $f_a(\theta, a_1)$ depends on θ. Given the simplification that $f(\theta, a_t) = \theta + ka_t$, however, $a_1^*(\theta)$ is independent of θ, though the main result of this paper is robust to the functional form of $f(\theta, a_t)$. Thus, I let $a_1^*(\theta) = a_1^*$ and $a_1^e(\theta) = a_1^e$. One may expect the wage wedge $[w_2(1) - w_2(0)]$ is positive. This is true because better date-1 performance implies that hidden productivity is larger, which means expected date-2 performance would be greater. The following lemma provides a formal account.

Lemma 3.1. *The marginal incentive in (3.9) is positive, independent of* θ*, and equal to*

$$\delta k \frac{\textsc{Var}^m(\theta)}{\textsc{Var}^m(y_1)} \tag{3.10}$$

where $\textsc{Var}^m$ *is market's assessed variance conditional on* $\mathbf{a}_t^e$*.*

Proof: First of all, the marginal incentive is positive since variances are always nonnegative and θ is nondegenerate random variable in market's point of view. Independence with respect to θ is obvious. Next, I shall prove that $[w_2(1) - w_2(0)]$ equals the ratio of variances. At first, the wage is given by

$$\begin{aligned} w_2(y_1) &= \mathbb{E}^m[y_2|y_1] = \mathbb{E}^m[\mathbb{E}^m[y_2|\theta, y_1]|y_1] \text{ by iterated expectation} \\ &= \mathbb{E}^m[1 \cdot f(\theta, a_2^e) + 0 \cdot \{1 - f(\theta, a_2^e)\}|y_1] \\ &= \underbrace{f(h, a_2^e)}_{=h} \mathbb{P}^m(h|y_1) + \underbrace{f(l, a_2^e)}_{=0} \{1 - \mathbb{P}^m(h|y_1)\} \end{aligned} \tag{3.11}$$

where $\mathbb{P}^m(h|y_1)$ is market's posterior of $\theta = h$ given y_1 and $\mathbf{a}_t^e$. This posterior is given by

$$\mathbb{P}^m(h|y_1 = 1) = \frac{\frac{1}{2} f(h, a_1^e)}{\frac{1}{2} f(h, a_1^e) + \frac{1}{2} f(l, a_1^e)}$$

$$\mathbb{P}^m(h|y_1 = 0) = \frac{\frac{1}{2}[1 - f(h, a_1^e)]}{\frac{1}{2}[1 - f(h, a_1^e)] + \frac{1}{2}[1 - f(l, a_1^e)]}$$

Then, it follows

$$w_2(1) - w_2(0) = \frac{\frac{1}{4} f(h, a_2^e)[f(h, a_1^e) - f(l, a_1^e)]}{\frac{1}{2}[f(h, a_1^e) + f(l, a_1^e)][1 - \frac{1}{2}\{f(h, a_1^e) + f(l, a_1^e)\}]} \tag{3.12}$$

Note that the numerator equals $\frac{1}{4}h^2$, which equals $\text{VAR}^m(\theta)$ since

$$\text{VAR}^m(\theta) = \mathbb{E}^m[\theta^2] - \mathbb{E}^m[\theta]^2 = \frac{1}{2}h^2 - \frac{1}{4}h^2 = \frac{1}{4}h^2$$

The denominator equals $\text{VAR}^m(y_1)$ since

$$\begin{aligned} \text{VAR}^m(y_1) &= \mathbb{E}^m[y_1^2] - \mathbb{E}^m[y_1]^2 \\ &= \mathbb{E}^m[y_1](1 - \mathbb{E}^m[y_1]) \text{ since } y_1 \in \{0, 1\} \text{ implies } \mathbb{E}^m[y_1^2] = \mathbb{E}^m[y_1] \end{aligned}$$

where $\mathbb{E}^m[y_1] = \mathbb{E}^m[\mathbb{E}^m[y_1|\theta]] = \mathbb{E}^m[f(\theta, a_1^e)] = \frac{1}{2}\{f(h, a_1^e) + f(l, a_1^e)\}$. ■

Lemma 3.1 implies that date-1 marginal incentive is determined by a signal-to-noise ratio—the extent to which the data y_1 conveys information about hidden variable θ. Holmstrom (1999) found the same result under a simpler model in which neither market nor agent knows types. Thus, making θ private information does not affect the size of career-concerns' motive of work incentive under the current specification that expected output is linear in θ and a_t (i.e. $f(\theta, a_t) = \theta + ka_t$). In more general specification, work incentive depends on θ.

A key observation from (3.10) is as follows: if an agent's hidden ability is unknown to the labor market, the agent faces date-1 incentives in order to convince the market that she is of high ability.

Consider agent's payoff in a rational expectations equilibrium ($a_t^* = a_t^e$). First of all, the following lemma shows that a unique rational expectations equilibrium exists under reasonable restrictions on the effort cost function $c(a_t)$. Note that $w_2(y_1)$ depends on a_1^e. To highlight this dependence, I write $w_2(y_1) = w_2(y_1)(a_1^e)$.

Lemma 3.2. *Suppose (1)* $c'\left(\frac{1-h}{k}\right) > h\delta k$*. Then, there exists a rational expectations equilibrium* $a_1^* \in \left(0, \frac{1-h}{k}\right)$ *such that*

$$\delta k[w_2(1)(\hat{a}_1) - w_2(0)(\hat{a}_1)] = c'(\hat{a}_1) \text{ at } \hat{a}_1 = a_1^*$$

This rational expectations equilibrium is unique if (2) $c'''(\cdot) \geq 0$*.*

Proof: Let $v(a_1^e) \equiv w_2(1)(a_1^e) - w_2(0)(a_1^e) > 0$. And, let $g(\cdot) \equiv \frac{c'(\cdot)}{\delta k}$. Then, (*) there exists a unique $a_1 \in \left(0, \frac{1-h}{k}\right)$ for each given a_1^e such that $g(a_1) = v(a_1^e)$ since $c'(0) = 0$, $c'\left(\frac{1-h}{k}\right) > h\delta k$ and $h > v(a_1^e)$. Then, $a_1 = g^{-1}(v(a_1^e))$. Let $\Psi(a_1^e) \equiv g^{-1}(v(a_1^e))$. $\Psi(a_1^e)$ is a continuous function from a closed interval $\left[0, \frac{1-h}{k}\right]$ to the same interval. Then, Brouwer fixed point theorem and (*) imply that there exists an $a_1^* \in \left(0, \frac{1-h}{k}\right)$ such that $a_1^* = \Psi(a_1^*)$. That is, there exists a rational expectations equilibrium a_1^*. Note that $v(a_1^e)$ is positive, convex and U-shaped function. Then, (1), (2) and $c'(0) = 0$ imply the uniqueness. ■

In the unique rational expectations equilibrium a_1^*, date-1 wage equals

$$w_1 = \mathbb{E}^m[y_1] = \mathbb{E}^m[\mathbb{E}^m[y_1|\theta]] = \mathbb{E}^m[f(\theta, a_1^*)] = \frac{1}{2}h + ka_1^*$$

And, by (3.11), a type-θ agent's expectation of date-2 wage equals

$$\mathbb{E}^\theta[w_2(y_1)] = h\mathbb{E}^\theta[\mathbb{P}^m(h|y_1)]$$

Then, a type-θ agent's payoff equals

$$U_P(\theta) \equiv \frac{1}{2}h + ka_1^* - c(a_1^*) + \delta h \mathbb{E}^\theta[\mathbb{P}^m(h|y_1)] \tag{3.13}$$

Note that low-type agents will always participate in the labor market, as their reservation payoff u^l is zero. However, high-type agents will participate if and only if

Assumption 3.3. $\frac{1}{2}h + ka_1^* - c(a_1^*) + \delta h \mathbb{E}^h[\mathbb{P}^m(h|y_1)] \geq (1+\delta)u^h$

To see when high-type agents profitably participate, suppose for the moment $k = 0$ and hence $a_1^* = 0$ by (3.10). Notice that $\mathbb{E}^\theta[\mathbb{P}^m(h|y_1)] \in (0,1)$ is type-θ agents' expectation of market's posterior given y_1. High-type agents are more optimistic about the future than low-type agents, that is, $\mathbb{E}^l[\mathbb{P}^m(h|y_1)] < \frac{1}{2} < \mathbb{E}^h[\mathbb{P}^m(h|y_1)]$. This is related to market's learning effect. The labor market updates its expectation of date-2 output by observing date-1 output. High-type agents then expect more income since date-1 output is (stochastically) higher if the given agent is of high-type rather than low-type. If the learning effect is little (i.e. $\left|\mathbb{E}^h[\mathbb{P}^m(h|y_1)] - \frac{1}{2}\right|$ is little), Assumption 3.1 implies Assumption 3.3 is violated. If instead the learning effect is large enough, Assumption 3.3 could be satisfied. Even if the learning effect is little, if $k > 0$ and the surplus $ka_1^* - c(a_1^*)$ generated by career-concerns is large enough, then Assumption 3.3 is satisfied.

As long as high-type agents find it optimal to participate (i.e. Assumption 3.3 holds), the total surplus in this pooling equilibrium is $\frac{1}{2}[U_P(h)+U_P(l)]$, which equals

$$\frac{1}{2}(1+\delta)h + ka_1^* - c(a_1^*) \tag{3.14}$$

by the martingale property, which says that the expectation of posterior equals the prior.

3.2.3 Welfare comparison

Compare (3.14) to (3.4). One can see that the total surplus in the pooling equilibrium is greater than that in Riley separating equilibrium. That is, the use of education as a sorting device reduces total surplus. There are two reasons. First, education as a signaling device reveals hidden productivity. Thus, agents need not persuade their ability to the labor market, and thereby, their post-education work incentives motivated by career-concerns are eliminated. In consequence, the work-stage surplus $ka_1^* - c(a_1^*)$ is not realized. Second, education is wasteful and incurring cost $C(h, e(h))$.

The detrimental effect of education as a sorting device hinges on Assumption 3.3. If this assumption is not satisfied, only low-type agents participate in the labor market. The market then pays zero to the participants. Thus, the total surplus in the pooling equilibrium equals

$$\frac{1}{2}(1 + \delta)u^h$$

Thus, it goes back to the standard result that the use of education as a signaling device improves total surplus. Then, when Assumption 3.3 is more likely to be satisfied? It depends on h and u^h. The degree of the lemon market problem is measured by $|h - u^h|$ since adverse selection discourages high-type agents' participation in the labor market. Note that Assumption 3.3 holds if h is large or u^h is little since in this case high-type will participate even if the work-stage surplus is relatively low. Thus, Assumption 3.3 is satisfied if the degree of the lemon market problem is severe, which is of high interest, whereas it is violated if the lemon market problem is not very important, which is of little interest. In the standard Spence model in which post-education work-stage is ignored, education as a sorting device is more beneficial if the lemon market problem is more severe. But in this very case, education as a sorting device is more detrimental if post-education work-stage is explicitly considered.

Summarizing the analysis, I have the following first main result:

Proposition 3.1. *Suppose there is the lemon market problem (i.e. Assumption 3.1) and education is not prohibitively costly (i.e. Assumption 3.2). Then,*

(i) If there is no post-education work-stage, the use of education as a sorting device increases total surplus.

Suppose there is post-education work-stage.

(ii) If high-type agents profitably participate in the labor market in the most efficient pooling equilibrium (i.e. Assumption 3.3), which is the case when the lemon market problem is severe, then education as a sorting device decreases total surplus.

(iii) If high-type agents cannot profitably participate in the most efficient pooling equilibrium (i.e. Assumption 3.3 is violated), which is the case when the lemon market problem is unimportant, then education as a sorting device increases total surplus.

In the following, I consider two extensions. In the first extension, I consider multiple types. Productivity is determined by various hidden factors. For instance, intelligence and fitness-to-work are two important factors that contribute to overall productivity. Let θ_1 denote intelligence, which agent knows but market does not observe. Marginal cost of education $C_e(\theta_1, e)$ decreases in θ_1. Let θ_2 denote fitness-to-work, which neither agent nor market observes. Let $\theta \equiv g(\theta_1, \theta_2)$ be agent's productivity, where g is increasing in each argument. In this case, education can reveal only θ_1. Lemma 3.1

then implies that an agent's post-education marginal incentive in separating equilibria equals

$$\delta k \frac{\text{Var}^m(\theta|\theta_1)}{\text{Var}^m(y_1)}$$

while the marginal incentive in pooling equilibria equals

$$\delta k \frac{\text{Var}^m(\theta)}{\text{Var}^m(y_1)}$$

Unless θ_1 and θ_2 are independent, marginal incentive is greater in pooling equilibria. That is, the more information education reveals, the more likely it is detrimental to welfare.

In the second extension, I assume that education has both human capital enhancing effect and sorting effect. This situation can be modeled in the following way: the overall productivity θ' equals $\theta + \alpha e$ where $\alpha > 0$. In this case, the welfare implication of education is ambiguous. On the one hand, it improves efficiency by raising productivity. On the other hand, it reduces efficiency by discouraging post-education work incentives. This extension suggests that governmental education support (which aims at realizing human capital enhancement) should be universal rather than selective so that sorting effect is suppressed.

Proposition 3.1 implies that uncertainty in types is beneficial for social welfare. It seems at odd at a first glance since one might expect that the uncertainty, as a source of market failure, reduces social welfare. But it is consistent with the general theory of second-best suggested by Lipsey and Lancaster (1956): if there is an existing source of market failure (so that the economy is in a second-best outcome), an additional source of market failure could either increase or decrease social welfare. (See Milgrom and

Roberts (1982) and Kim (2004) for other examples of the general theory). However, Proposition 3.1 is not a simple corollary of the general theory. In fact, the general theory is too general to describe an underlying mechanism and characterize the conditions that make the mechanism.

Proposition 3.1 is related to the literature of information disclosure. In a dynamic tournament setting, Ederer (2010) shows that disclosing interim performance could reduce incentives due to the tradeoff between evaluation and motivation effects. In the presence of career-concerns and relational contracting, Mukherjee (2008) show that if the current employer discloses workers' performance to the labor market, it increases career-concerns' motive of incentives while reduces the effectiveness of relational contracting. Proposition 3.1 is however different from these papers in that (1) it examines the effect of disclosing types rather than performance on incentives and (2) these papers consider only post-education workplace behaviors while my paper considers both education and post-education behaviors.

3.2.4 Intuitive Criterion

There are infinitely many separating and pooling equilibria in standard job market signaling model (where there is no post-education work-stage). However, Cho and Kreps (1987) show that only Riley separating equilibrium (and no pooling equilibria) satisfies their Intuitive Criterion, which is nowadays a standard equilibrium refinement criterion, and hence reasonable. Recently, Alos-Ferrer and Prat (2012) show that some pooling equilibria can satisfy the Intuitive Criterion if there is learning by the market (or employers). If there is post-education work-stage as in this paper, pooling equilibria become more likely consistent with the Intuitive Criterion. These points are elaborated in the following.

Separating equilibrium

Consider at first standard job market signaling model. Let (e^l, e^h) be a non-Riley separating equilibrium such that $e^l = 0$ and $e^h > e(h)$, where $e(h)$ is given by (3.3). (Note that e^l cannot be positive since otherwise low-type profitably deviates to zero education). Then, if an agent chooses e' such that $e(h) < e' < e^h$, the market should *not* believe that the agent is of high-type since otherwise high-type profitably deviates to e'. However, such belief is not *intuitive* for the following reason. If low-type deviates to e', she is worse off even in the best scenario in which the market believes her as high-type and hence pays h rather than l. If high-type deviates to e', he is better off in the best scenario. That is, low-type never deviates but high-type may deviate. Thus, the market should believe that those who deviate to e' are of high-type. Therefore, if only *intuitive* belief is allowed, as the Intuitive Criterion requires, (e^l, e^h) is no longer a separating equilibrium. Nothing changes in this argument even if one introduces post-education work-stage, as there is no uncertainty in types.

Pooling equilibrium

Consider standard job market signaling model. Let e_P be a pooling equilibrium. Then, there are e_P^θ such that

$$\mathbb{E}^m[\theta](1+\delta) - C(\theta, e_P) = h(1+\delta) - C(\theta, e_P^\theta)$$

It follows then $e_P < e_P^l < e_P^h$. For e_P being a pooling equilibrium, the market should *not* believe that the agent who deviates to $e' \in (e_P^l, e_P^h)$ is of high-type, since otherwise high-type profitably deviates to e'. But this belief is not *intuitive*. By deviating to e', low-type is worse off even in the best scenario (i.e. the market believes her as high-type) while high-type is better off in the best scenario. Thus, only high-type may deviate to e', and hence, the market should believe that those who choose e' are of high-type. If only *intuitive* belief is allowed, e_P is no longer a pooling equilibrium.

However, this is not the case if we add post-education learning by market (or employers). After education choice, output signal y_t is realized. (To focus on the learning effect, I assume for now $k = 0$ so that agents do not exert any work effort). Observing y_1, the labor market learns (partially) about hidden type and then adjusts its expectation of date-2 output $\mathbb{E}^m[\theta|y_1]$. High-type agents know that their date-1 output will be (stochastically) better than low-type agents. Thus, even if types are not revealed in pooling equilibria, high-type agents expect larger income (i.e. $\mathbb{E}^h[\mathbb{E}^m[\theta|y_1]] > \mathbb{E}^l[\mathbb{E}^m[\theta|y_1]]$). Then, high-type's gain from separating himself (by taking higher education than e_P) from low-type is lower than it is when there is no post-education learning. To see this, note that e_P^θ is determined via

$$\mathbb{E}^m[\theta] + \delta\mathbb{E}^\theta[\mathbb{E}^m[\theta|y_1]] - C(\theta, e_P) = h(1+\delta) - C(\theta, e_P^\theta)$$

If this learning effect (or more precisely the effect of market's learning on the difference between high- and low-type agents' expected date-2 income $\mathbb{E}^h[\mathbb{E}^m[\theta|y_1]] - \mathbb{E}^l[\mathbb{E}^m[\theta|y_1]]$) is large enough, it follows that $e_P^l > e_P^h$. Then, there is no education level e' such that low-type is worse off even in the best scenario while high-type is better off in the best scenario. Thus, this pooling equilibrium satisfies the Intuitive Criterion. (See Alos-Ferrer and Prat (2012) for more detail).

If one also considers post-education working (i.e. $k > 0$), then pooling equilibria becomes more robust with respect to the Intuitive Criterion. By separating himself from low-type, high-type loses the work-stage surplus $ka_1^* - c(a_1^*)$, which realizes only in pooling equilibria. If this surplus is large enough, high-type is worse off by separating himself from low-type (by taking higher education than e_P) even in the best scenario. That is, for any $e > e_P$, it follows that

$$\mathbb{E}^m[\theta] + \delta\mathbb{E}^h[\mathbb{E}^m[\theta|y_1]] + ka_1^* - c(a_1^*) - C(h, e_P) > h(1 + \delta) - C(h, e)$$

The pooling equilibrium e_P then satisfies the Intuitive Criterion.

3.3 Model II: Education-Career Path Selection

In the previous section, I compare two distinct frames, one in which education is used as a sorting device (i.e. Riley separating equilibrium) and the other in which education is not used as a sorting device (i.e. Pooling equilibrium with zero education level). Perhaps, the society or government chooses a frame. But a frame is given to all agents.

In this section, agents can choose one from two types of occupation. In the first type of occupation, compensation is primarily based on achievement. For instance, entrepreneurs or traders in investment banks face very volatile compensation with respect to their performance. In the second type of occupation, pays are mainly determined by education. For instance, in think tanks, research fellows are mostly Ph.D. holders and their pays are relatively less volatile with respect to (research) performance in comparison to the first type of occupation. I shall model the first type as the pooling equilibrium with zero (or minimum) education and the second type as Riley separating equilibrium. I call the first type as pooling occupation (or performance-sensitive job) and the second type as separating occupation (or education-sensitive job).

Let λ denote the proportion of high-type agents among those who participate in the pooling occupation. If all high-type and all low-type agents participate in the pooling occupation, $\lambda = \frac{1}{2}$. If half of high-type agents and all of low-type agents participate, $\lambda = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{2}} = \frac{1}{3}$. Note that λ will be endogenously determined in equilibrium. Let $\mathbb{E}^m_\lambda$ denote market's expectation conditional on $\mathbf{a}^e_t$ and λ.

If participates in the separating occupation, a type-θ agent gets the following payoff

$$U^\theta_S \equiv \theta(1+\delta) - C(\theta, \hat{e}(\theta))$$

where $\hat{e}(\theta)$ denotes the equilibrium education level such that $\hat{e}(l) = 0$ and $\hat{e}(h) > 0$ is the minimum level low type cannot profitably mimc. Note that $\hat{e}(\theta)$ may be different from $e(\theta)$ defined in (3.3) since low-type agents have the option to choose the pooling occupation, which means their reservation payoff may be greater than zero. More detail will be presented momentarily.

Instead, if participates in the pooling occupation, an agent gets either

$$U_P(1) \equiv \mathbb{E}^m_\lambda[\theta] + ka^*_1 - c(a^*_1) + \delta w_2(1)$$

if $y_1 = 1$ is realized in the end of date-1 or

$$U_P(0) \equiv \mathbb{E}^m_\lambda[\theta] + ka^*_1 - c(a^*_1) + \delta w_2(0)$$

if $y_1 = 0$ is realized in the end of date-1.

That is, participating in the pooling occupation generates a random payoff $U_P(y_1)$. For low-type agents, this randomness does not matter for their choice of occupation. This is because they always prefer pooling to separating occupation no matter what y_1 is realized (i.e. $U_S^l = 0 < U_P(0) < U_P(1)$), as $\mathbb{E}_\lambda^m[\theta] \in [0, h]$, $\delta w_2(y_1) \in (0, h)$ for $y_1 = 0, 1$ and the work-stage surplus $ka_1^* - c(a_1^*)$ is positive. Next, consider the payoffs of high-type agents. If the work-stage surplus is either too high (i.e. $U_S^h < U_P(0) < U_P(1)$) or too low (i.e. $U_P(0) < U_P(1) < U_S^h$), then the randomness in $U_P(y_1)$ still does not affect their occupation choice.

However, if the work-stage surplus is moderate (i.e. $U_P(0) < U_S^h < U_P(1)$), this randomness becomes income risk. Suppose that high-type agents are heterogenous in terms of their risk-aversion r, which is uniformly distributed on $[0, \bar{r}]$. Then, switching separating to pooling occupation incurs to a r-type of high-type agent the following disutility:

$$\frac{1}{2}r\{U_P(1) - U_P(0)\} = \frac{1}{2}r\delta\{w_2(1) - w_2(0)\}$$

where $\frac{1}{2}$ is a rescaling constant. Thus, a r-type of high-type agents prefers the pooling occupation to the separating occupation if and only if

$$\mathbb{E}^h\left[U_P(y_1)\right] - \frac{1}{2}r\delta\{w_2(1) - w_2(0)\} \geq U_S^h \tag{3.15}$$

From (3.15) I have the following implications. First, if the mean return $\mathbb{E}^h\left[U_P(y_1)\right]$ from the pooling occupation is lower than the certainty return U_S^h from the separating occupation, then no high-type agents choose the pooling occupation. Second, if the pooling occupation provides greater mean return, those with low risk-aversion choose the pooling occupation while those with high risk-aversion choose the separating occupation. These two implications explain why some people choose little education and pursue high-risk and high-return jobs, whereas others choose decent degree of education and pursue low-risk and medium-return jobs. Many smart but risk-tolerate people choose to take only college level of education and start their career early as traders in investment banks or as entrepreneur in startups, which promises high-return with high-risk. Other smart but risk-averse people take Ph.D. degrees and then work for government or think-tanks, which gives medium-return with relatively low income risk.

This section illustrates a new theory of education. Traditional theories consider human capital effect or signaling effect of education. The current analysis suggests income-risk hedging effect of education. If the labor market observes a decent degree of education, it will pay for it assuming that the agent is of high quality. Accordingly, pay is not much dependent on stochastic performance. Agents with high risk-aversion then invest on education in order to hedge against future income risk.

3.3.1 Equilibrium characterization

The current analysis is somewhat preliminary, as endogenous prior λ, education level $\hat{e}(\theta)$ and other equilibrium conditions are not explicitly specified. In the following, I characterize equilibrium conditions and consider a numerical example consistent with the conditions.

I confine my attention to an interesting case in which some people participate in the pooling occupation and others do in the separating occupation, as people choose diverse jobs in the real world. (Other less interesting cases are such that everyone participates only in the pooling occupation or only in the separating occupation). I first provide conditions that characterize such an equilibrium $(\hat{e}(h), w_2(y_1), a_1^e, a_1^*, U_P(y_1), U_S^h, r^h, \lambda)$ and then explain these conditions

(1) $\hat{e}(h) = \min\left\{e \geq 0 : (1+\delta)h - C(l,e) \leq \mathbb{E}^l[U_P(y_1)]\right\}$

(2) $w_2(1) = h\frac{\frac{1}{2}(h+ka_1^e)}{\frac{1}{2}h+ka_1^e}$ and $w_2(0) = h\frac{\frac{1}{2}(1-h-ka_1^e)}{1-\frac{1}{2}h-ka_1^e}$

(3) $a_1 = a_1^* \in (0, \bar{a})$ solves $\delta k[w_2(1) - w_2(0)] = c'(a_1)$ for a given a_1^e [1]

(4) $a_1^* = a_1^e$

(5) $U_P(y_1) = \lambda h + ka_1^* - c(a_1^*) + \delta w_2(y_1)$

(6) $U_S^h = h(1+\delta) - C(h, \hat{e}(h))$

(7) $U_P(0) < U_S^h < U_P(1)$

(8) $\mathbb{E}^h[U_P(y_1)] - \frac{1}{2}\bar{r}\,[w_2(1) - w_2(0)] < U_S^h < \mathbb{E}^h[U_P(y_1)]$

(9) $r = r^h \in (0, \bar{r})$ solves $\mathbb{E}^h[U_P(y_1)] - \frac{1}{2}r\,[w_2(1) - w_2(0)] = U_S^h$

(10) $\lambda = \frac{\frac{1}{2}\frac{r^h}{\bar{r}}}{\frac{1}{2}\frac{r^h}{\bar{r}}+\frac{1}{2}} = \frac{r^h}{r^h+\bar{r}}$

(1) characterizes the minimum level of education $\hat{e}(h)$, which discourages low-type agents' mimicking high-type in the separating occupation. Note that low-type agents get a positive payoff in the pooling occupation (due to work-stage surplus and pooling of wage). Thus, low-type agents find mimicking unprofitable if the resulting payoff is less than what they get in the pooling occupation. (2) shows how wage $w_1(y_1)$ is formed. In (3), equilibrium effort a_1^* is characterized given market's anticipated effort a_1^e. And, as in (4), they

[1]I assume that an interior solution exists, that is, $a_1^* < \bar{a}$.

must be the same in rational expectations equilibrium. High-type agents' certainty payoff U_S^h in the separating occupation is given in (5) and each type's output-contingent payoff $U_P(y_1)$ in the pooling occupation is given in (6). (7) and (8) restrict parameters so that we are in the following interesting equilibrium: high-type agents face income risk if they switch from the separating to the pooling occupation; the pooling occupation promises greater mean payoff than the separating occupation, though this greater (but risky) payoff is not enough appealing to those with the highest degree of risk-aversion. (9) characterizes the marginal high-type agents who are indifferent between pooling and separating jobs. (10) characterizes the endogenous prior that participants in the pooling occupation are of high-type.

The above equilibrium is involved with so many variables and conditions that closed-form characterization is usually infeasible. In the following, I provide a numerical example that satisfies all the conditions above and hence constitutes an equilibrium.

A numerical equilibrium. Consider the following cost functions: $C(\theta, e) = \frac{1}{2(4+\theta)}e^2$ and $c(a_t) = \frac{1}{3}a_t^3$. Consider the following numerical values: $k = .73$, $\delta = .95$, $h = .45$, $\bar{a} = .4$ and $\bar{r} = 1$. Then, I have the following numerical equilibrium: $\hat{e}(h) = 1.777$, $a_1^* = .352$, $w_2(1) = .244$, $w_2(0) = .065$, $U_P(1) = .609$, $U_P(0) = .439$, $U_S^h = .523$, $r_h = .429$ and $\lambda = .3$.

The numerical equilibrium above (and also in general) provides three noteworthy observations. First, low-type agents are better off switching from the separating to the pooling occupation, that is, $U_P(1) > U_P(0) > 0$. Therefore, only high-type agents participate in the separating occupation. Second, high-type agents participating in the separating occupation must take a positive level of education (i.e. $\hat{e}(h) > 0$) even if there are no low-type agents in the same occupation. This implies that education is not used to sort types *within* the same labor market. Rather, education is used a sort of entry barrier. Note that low-type agents would be better off switching from the pooling to the separating occupation if they were not required to take any positive education, as $U_P(0) < U_P(1) < (1+\delta)h = .8775$. In equilibrium, however, low-type agents need to take a sufficient amount of costly education $\hat{e}(h) > 0$, which deters their entry to the separating occupation. Thus, education as an entry barrier is used to allocate high-type agents to education-sensitive jobs and low-type agents to performance-sensitive jobs. Third, one can reconfirm that education is used as an income-risk hedging device. Less risk-averse high-type agents (whose risk aversion r is less than $r_h = .429$) undertake the income-risk in order to get higher mean payoff (i.e. $\mathbb{E}^h[U_P(y_1)] = .559 > .523 = U_S^h$). In contrast, more risk-averse hype-type agents (whose risk aversion r is greater than $r_h = .429$) takes costly education $\hat{e}(h)$ in order to hedge the income-risk and gets lower but safe payoff.

Bibliography

Acharya, Viral, Itamar Drechsler, and Philipp Schnabl (2014), "A pyrrhic victory? bank bailouts and sovereign credit risk." *Journal of Finance*, 69, 2689–2739.

Acharya, Viral V. and Tanju Yorulmazer (2008), "Cash-in-the-market pricing and optimal resolution of bank failures." *Review of Financial Studies*, 21, 2705–2742.

Adams, Renee and Daniel Ferreira (2007), "A theory of friendly boards." *Journal of Finance*, 62, 217–250.

Aghion, P., P. Bolton, and S. Fries (1999), "Optimal design of bank bailouts: the case of transition economies." *Journal of Institutional and Theoretical Economics*, 155, 51–79.

Alos-Ferrer, Carlos and Julien Prat (2012), "Job market signaling and employer learning." *Journal of Economic Theory*, 147, 1787–1817.

Baker, George, Robert Gibbons, and Kevin Murphy (1994), "Subjective performance measures in optimal incentive contracts." *The Quarterly Journal of Economics*, 109, pp. 1125–1156.

Benabou, Roland and Jean Tirole (2013), "Bonus culture: Competitive pay, screening, and multitasking." NBER Working Papers 18936, National Bureau of Economic Research, Inc.

Bernardo, A., E. Talley, and I. Welch (2011), "A model of optimal government bailouts." *Berkeley Olin Program in Law & Economics, Working Paper Series.*

Bernheim, Douglas and Michael Whinston (1998), "Incomplete contracts and strategic ambiguity." *American Economic Review*, 88, pp. 902–932.

Boot, A. and A Thakor (1993), "Self-interested bank regulation." *American Economic Review AEA Papers and Proceedings*, 83, 206–212.

Brown, Craig O. and I. Serdar Dinc (2005), "The politics of bank failures: Evidence from emerging markets." *Quarterly Journal of Economics*, 120, 1413–1444.

Brunnermeier, Markus, Andrew Crockett, Charles Goodhart, Avinash D. Persaud, and Hyun Song Shin (2009), "Geneva reports on the world economy 11: The fundamental principles of financial regulation." *ICMB, CEPR*, 1041–1063.

Chari, Varadarajan V. and Patrick J. Kehoe (2010), "Bailouts, time-inconsistency, and optimal regulation." *Mimeo Minneapolis Fed and Princeton.*

Cho, In-Koo and David M Kreps (1987), "Signaling games and stable equilibria." *Quarterly Journal of Economics*, 179–221.

Conlon, John (2009), "Two new conditions supporting the first-order approach to multisignal principal-agent problems." *Econometrica*, 77, 249–278.

Cordella, T. and E. Yeyati (2003), "Bank bailouts: Moral hazard vs. value effect." *Journal of Financial Intermediation*, 12, 300–330.

Dell'Ariccia, Giovanni and Lev Ratnovski (2013), "Bailouts, contagion, and moral hazard." *Working paper*.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999), "The economics of career concerns, part i: Comparing information structures." *Review of Economic Studies*, 66, 183–198.

Ederer, Florian (2010), "Feedback and motivation in dynamic tournaments." *Journal of Economics and Management Strategy*, 19, 733–769.

Freixas, X (1999), "Optimal bail-out, conditionality and constructive ambiguity." *Financial Market Group Discussion Paper, London School of Economics*, 237.

Fuchs, William (2007), "Contracting with repeated moral hazard and private evaluations." *American Economic Review*, 97, pp. 1432–1448.

Gibbons, Robert (2010), "Inside organizations: Priging, politics, and path dependence." *Annual Review of Economics*, 2, 337–365.

Gibbons, Robert and Kevin J. Murphy (1990), "Relative performance evaluation for chief executive officers." *Industrial and Labor Relations Review*, 43, 30–51.

Goodhart, C and H Huang (1999), "A model of the lender of last resort." *IMP Working Paper*.

Grossman, Sanford and Oliver Hart (1983), "An analysis of the principal-agent problem." *Econometrica*, 7–45.

Heckman, James, Carolyn Heinrich, and Jeff Smith (1997), "Assessing the performance of performance standards in public bureaucracies." *American Economic Review*, 87, 389–395.

Hochberg, Yael, Alexander Ljungqvist, and Yang Lu (2007), "Whom you know matters: Venture capital networks and investment performance." *Journal of Finance*, 62, 251–301.

Holmstrom, Bengt (1979), "Moral hazard and observability." *Bell Journal of Economics*, 10, 74–91.

Holmstrom, Bengt (1999), "Managerial incentive problems: A dynamic perspective." *Review of Economic Studies*, 66, 169–182.

Holmstrom, Bengt and Paul Milgrom (1991), "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law and Economics and Organization*, 7, 24.

Hwang, Sunjoo (2014), "Relational contracts and the value of information." *Working paper*.

Innes, Robert (1990), "Limited liability and incentive contracting with ex-ante action choices." *Journal of Economic Theory*, 52, 45–67.

Jewitt, Ian (1988), "Justifying the first-order approach to principal-agent problems." *Econometrica*, 1177–1190.

Jewitt, Ian, Ohad Kadan, and Jeroen Swinkels (2008), "Moral hazard with bounded payments." *Journal of Economic Theory*, 143, 59–82.

Kandori, Michihiro (2002), "Introduction to repeated games with private monitoring." *Journal of Economic Theory*, 102, 1 – 15.

Kim, Jaehong (2004), "Efficiency of entry regulation under incomplete information." *mimeo.*

Kim, Sonku (1995), "Efficiency of an information system in an agency model." *Econometrica*, 89–102.

Kornai, J., E. Maskin, and G. Roland (2003), "Understanding the soft budget constraint." *Journal of Economic Literature*, 41, 1095–1136.

Kornai, Janos (1979), "Resource-constrained versus demand-constrained systems." *Econometrica*, 47, 801–819.

Kvaloy, Ola and Trond Olsen (2014), "Teams and tournaments in relational contracts." CESifo Working Paper Series 4783, CESifo Group Munich.

Levin, Jonathan (2003), "Relational incentive contracts." *American Economic Review*, 93, 835–857.

Lewellen, Stefan (2013), "Executive compensation and peer effects." *Working Paper.*

Lipsey, R. and Kelvin Lancaster (1956), "The general theory of second best." *Review of Economic Studies*, 24, 11–32.

MacLeod, Bentley (2003), "Optimal contracting with subjective evaluation." *American Economic Review*, 93, 216–240.

Milgrom, Paul and John Roberts (1982), "Limit pricing and entry under incomplete information: an equilibrium analysis." *Econometrica*, 50, 443–459.

Milgrom, Paul and Rober Weber (1982), "A theory of auctions and competitive bidding." *Econometrica*, 50, 1089–1122.

Mirrlees, James (1976), "The optimal structure of incentives and authority within an organization." *Bell Journal of Economics*, 7, pp. 105–131.

Mirrlees, James (1999), "The theory of moral hazard and unobservable behaviour: Part i." *Review of Economic Studies*, 66, 3–21.

Mukerji, Sujoy (1998), "Ambiguity aversion and incompleteness of contractual form." *American Economic Review*, 88, 1207–31.

Mukherjee, Arijit (2008), "Sustaining implicit contracts when agents have career concerns:the role of information disclosure." *RAND Journal of Economics*, 39, 469–490.

Murphy, Kevin J. (1999), "Chapter 38 executive compensation." *Handbook of Labor Economics*, Volume 3, Part B, 2485–2563.

Murphy, Kevin J. (2009), "Compensation structure and systemic risk." *Testimony in front of the Committee of Financial Services, United States of House Representatives*.

Myers, Stewart and Raghuram Rajan (1998), "The paradox of liquidity." *Quarterly Journal of Economics*, 113, 733–771.

Pazarbasioglu, Ceyla, Jianping Zhou, Vanessa Le Lesle, and Michael Moore (2011), "Contingent capital: Economic rationale and design features." *IMF Staff Discussion Note*.

Pourciau, Susan (1993), "Earnings management and nonroutine executive changes." *Journal of Accounting and Economics*, 16, 317–336.

Rogerson, William (1985), "The first-order approach to principal-agent problems." *Econometrica*, 53, 1357–1367.

Rothschild, Michael and Joseph Stiglitz (1970), "Increasing risk: I. a definition." *Journal of Economic Theory*, 2, 225–243.

Schaffer, M. (1989), "The credible-commitment problem in the center-enterprise relationship." *Journal of Comparative Economics*, 13, 359–382.

Schmidt, Klaus M. (1996), "The costs and benefits of privatization: An incomplete contracts approach." *Journal of Law, Economics, and Organization*, 12, 1–24.

Segal, I.R. (1998), "Monopoly and soft budget constraint." *RAND Journal of Economics*, 29, 596–609.

Shin, J. and H. Chang (2003), *Restructuring Korea Inc*, volume 42. Routledge.

Sinclair-Desgagne, Bernard (1994), "The first-order approach to multi-signal principal-agent problems." *Econometrica*, 62, 459–66.

Sinclair-Desgagne, Bernard (2009), "Ancillary statistics in principal-agent models." *Econometrica*, 77, 279–281.

Spence, Michael (1973), "Job market signaling." *Quarterly Journal of Economics*, 87, 355–374.

Tirole, Jean (2010), *The Theory of Corporate Finance*. Princeton University Press.

Weiss, Andrew (1995), "Human capital vs. signaling explanations of wages." *Journal of Economic Perspectives*, 9, 133–154.